ADVANCED CALCULUS

ADVANCED CALCULUS

A SEQUEL TO

AN ELEMENTARY TREATISE ON THE CALCULUS

BY THE LATE

GEORGE A. GIBSON

M.A., LL.D., F.R.S.E.

EMERITUS PROFESSOR OF MATHEMATICS, UNIVERSITY OF GLASGOW

MACMILLAN & CO LTD

First Edition 1931
Reprinted 1944, 1948, 1954, 1958

MACMILLAN AND COMPANY LIMITED
Bombay Calcutta Madras

CONTENTS

CHAPTER I

REAL NUMBERS. SECTIONS

CHAPTER II

SETS. SEQUENCES. LIMITING POINTS. LIMITS

CHAPTER III

FUNCTIONS OF ONE VARIABLE. DERIVATIVES. DIFFERENTIALS

CHAPTER IV

FUNCTIONS OF SEVERAL VARIABLES. DERIVATIVES. DIFFERENTIALS. CHANGE OF VARIABLES

CHAPTER V

IMPLICIT FUNCTIONS. JACOBIANS

CHAPTER VI

INFINITE SERIES. COMPLEX FUNCTIONS OF A REAL VARIABLE

CHAPTER VII

SUBSTITUTION OF A SERIES IN A SERIES. REVERSION OF SERIES. LAGRANGE'S EXPANSION. MAXIMA AND MINIMA OF FUNCTIONS OF SEVERAL VARIABLES

CHAPTER VIII

INFINITE PRODUCTS. PRODUCTS AND SERIES OF PARTIAL FRACTIONS FOR TRIGONOMETRIC FUNCTIONS. GAMMA FUNCTIONS

CHAPTER IX

INTEGRATION OF BOUNDED FUNCTIONS

CHAPTER X

RECTIFICATION. CURVILINEAR INTEGRALS. AREAS. REPEATED AND DOUBLE INTEGRALS. VOLUMES. SURFACES

CHAPTER XI

MULTIPLE INTEGRALS. SURFACE INTEGRALS

CHAPTER XII

IMPROPER INTEGRALS

CHAPTER XIII

IMPROPER INTEGRALS: REPEATED AND DOUBLE INTEGRALS, FIELD OF INTEGRATION FINITE

CHAPTER XIV

DOUBLE INTEGRALS: RANGE INFINITE

CHAPTER XV

INTEGRATION OF SERIES. GAMMA FUNCTIONS

CHAPTER I

REAL NUMBERS. SECTIONS

1. The Continuous Variable. In § 9 of the *Elementary Treatise on the Calculus* a continuously varying number x is represented as the abscissa of a point P on an axis, and the assumption is made (§ 4) that, when an origin O and a unit segment OU have been chosen, there is a one-to-one correspondence between the numbers x and the points P. In § 148 the existence of a limit is discussed on the same assumption. Now the correspondence between numbers and points obviously fails if number is restricted to mean *rational number*, that is, positive or negative integer, or zero, or fraction (the quotient of an integer by an integer); for example, if OP is equal to the diagonal of a square of side OU there is no rational number that corresponds to P since the diagonal and the side of a square are incommensurable. It is customary to say that the abscissa of P is 1·41 or 1·414 or 1·4142, ... *approximately*; in other words, in order to make the one-to-one correspondence complete, it is assumed that there *is* a number corresponding to P, denoted by $\sqrt{2}$, and called an *irrational* number, and that 1·41, 1·414, 1·4142, ... are rational approximations to it.

The theory of irrational numbers is very much harder than that of rational numbers, and in books on elementary algebra the definition of an irrational number and the discussion of the laws of operation of such numbers are necessarily rather sketchy. It is desirable, however, that the advanced student should face the difficulties and see that a purely arithmetical theory of irrational numbers can be constructed; it will be found that the conception of the continuity of the straight line can be stated in such a way that the correspondence between "number" and "point" can be re-established.

Definitions of the irrational number have been given by Dedekind, Weierstrass, G. Cantor and Méray, and it is on one or other of these definitions that the treatment is now usually based. Of English textbooks to which the student may be referred for a discussion of the modern conception of number the most important is Hobson's *Theory of Functions of a Real Variable.* Good but less elaborate discussions will be found in Chrystal's *Algebra*, Part II, (2nd Ed.), pp. 97-109, Hardy's *Pure Mathematics* and Bromwich's *Infinite Series* (Appendix I). The sketch that follows is intended to call attention merely to the more important parts of the theory; as Chrystal remarks (*l.c.* p. 98), "the initial difficulties of the theory lie not in framing definitions, but in seeing where new definitions and where demonstrations are really necessary." The discussion now to be given is based on Dedekind's exposition in his tract *Stetigkeit und Irrationale Zahlen.*

2. Continuity of a Straight Line. Dedekind's definition of the irrational number will probably, at a first approach, seem rather strange to the student, and it may be helpful to sketch the geometric considerations that suggested his definition. The problem before him was that of finding a *mathematical* test of continuity such as would form the basis of mathematical deductions, and these geometric considerations may be stated briefly in the following way.

The following relations hold for points on a straight line:

(1) If A lies to the left of B and B to the left of C, then A lies to the left of C and B lies between A and C.

(2) If A and B are two different points on the line, there is an unlimited number of points between them.

(3) If A is a given point on the line it divides *all* the points of the line into two classes—the left or lower class (L say) and the right or upper class (U say). The class L contains all points to the left of A and the class U all points to the right of A, while the point A itself may be assigned either to L or to U.

This division of the points of the line is called a "section," and A is said to generate the section; the section is considered to be the same whether A is assigned to L or to U.

Now Dedekind considers that the characteristic property of

continuity may be expressed by the *converse* of (3), and he states it in the form: If all the points of a straight line are divided into two classes, L and U, so that every point in L lies to the left of every point in U then *there is one and only one point that generates the section.*

This statement is to be taken as an *Axiom*; it seems to be consistent with "common sense" conceptions of the continuity of the straight line and, when the irrational number has been defined in the purely arithmetical manner proposed by Dedekind, the expression $x \cdot OU$ for the length of a segment OP in terms of a standard unit segment OU is recovered.

The definition of the irrational number presupposes the knowledge of the rational numbers, and the student is advised to make a careful study of the early chapters of Chrystal's *Algebra* so that he may appreciate the gradual extension of the *meaning* of "number" and of the laws of operation on numbers. There may, for example, be three things in a group; the "natural" number 3 which specifies this characteristic of the group is not a number in the same sense as the "positive" number 3 although the same symbol is used for both and the word "number" is applied to both. The justification of the extension of the word "number" to positive, zero, negative or fractional numbers lies in the fact that the laws of operation on these numbers are consistent with those applicable to natural numbers and a similar justification holds in regard to "irrational" numbers.

Though no appeal is made to geometry the parallelism between the geometrical statements (1), (2) and (3) of this article and the corresponding arithmetical statements in the next article should be noted.

3. The Rational Number. In this article the word "number" means "rational number," and it is supposed that the laws of operation on such numbers are known; the expression "rational number" will only be employed when it seems to be desirable to emphasize the restriction.

The following relations hold for rational numbers:

(1) If $a<b$ and $b<c$ then $a<c$ and b is said to lie between a and c. Further, if a and b are two numbers one and only one of the following alternatives is true: $a>b$, or $a<b$, or $a=b$.

The system of rational numbers is therefore an *ordered* system; that is, just as a set* of points on a straight line, counted say from left to right, has a definite order, and, of any pair of different points, one always *precedes*, or *lies to the left of*, the other so, of any pair of different rational numbers, one always *precedes*, or *is less than*, the other while the numbers in any set of rational numbers have *inter se* a definite order, namely the order of magnitude.

The system of rational numbers will be denoted by R.

(2) If a and b are two different numbers, there is an unlimited number of numbers between them; in other words, if k is any positive integer, no matter how large, there are more than k numbers between a and b.

For if $a<b$ all the numbers

$$a+\frac{b-a}{n},\quad a+\frac{2(b-a)}{n},\quad a+\frac{3(b-a)}{n},\ \ldots,\ a+\frac{(n-1)(b-a)}{n},$$

where n is any integer greater than $(k+1)$, lie between a and b.

This property of the system R of rational numbers is expressed by saying that R is "dense" or "compact."

(3) If a is *any* number it separates all the numbers in R into two classes, a lower class L and an upper class U; the class L contains all numbers less than a and the class U all numbers greater than a, while a itself may be assigned either to L or to U.

If a is assigned to L it is the greatest number in L, and then U has no least number; if a is assigned to U it is the least number in U and then L has no greatest number. There cannot be *both* a greatest number a in L *and* a least number b in U because a and b would be different and all numbers between a and b would escape classification.

This separation of the numbers in R into two classes is called a *section* of R, and the number a is said to generate the section; a is either the greatest number in L or the least number in U.

The question at once arises: if *all* the numbers in R are separated into two classes L and U so that every number in L is less than every number in U, is there always a number in R that generates the section? Certainly not, as the following simple example shows.

* The word "set" means here simply "finite number.

It is easily proved * that there is no positive (rational) number whose square is 2, and therefore every positive number x is such that either $x^2 < 2$ or $x^2 > 2$. Now form two classes L and U of the numbers in R by assigning to L all the negative numbers, zero and the positive numbers whose square is less than 2, and to U all the other numbers in R, that is, all the positive numbers whose square is greater than 2. It will be shown that in this case L contains no greatest number and U no least.

Suppose a positive and $a^2 < 2$ so that $(2-a^2)$ is positive ; a greater number, $a+h$ say, can be found whose square is also less than 2. To see this, choose h so that $0 < h < 1$; then

$$(a+h)^2 < 2 \text{ if } 2ah + h^2 < 2 - a^2, \text{ a positive number.}$$

But $2ah + h^2 < 2ah + h$ since $0 < h < 1$ so that $(a+h)^2$ will be less than 2 if $(2a+1)h$ is less than $(2-a^2)$. It is possible to choose h so that $(2a+1)h$ will be less than $(2-a^2)$; for example, $(2a+1)h$ is less than $(2-a^2)$ if $h = (2-a^2)/(2a+2)$. It now follows that, whatever positive number be taken in L, there is always a greater number (in fact, an unlimited number of greater numbers) in L whose square is less than 2 so that L contains no greatest number.

In a similar way it may be shown that if a is positive and $a^2 > 2$ a positive number h may be found such that $a-h$ is positive and $(a-h)^2 > 2$; U therefore contains no least number.

The numbers in R have therefore been separated into two classes L and U such that every number in R occurs either in L or in U, and every number in L is less than every number in U, but there is in L no greatest number and in U no least so that the section is not generated by a number in R. The student will easily construct other examples of sections of R that are not generated by numbers in R. (See Exercises I.)

When a section is generated by a number in R that number may be said to *correspond* to the section, but when the section is not generated by a number in R there is no number with which to put the section in correspondence ; in the geometrical analogue of § 2 there is, on the assumption of Dedekind's Axiom, always one and only one point that

* See Exercises I, 1, 2.

generates a section. The arithmetical problem is now to extend the number system so that every section of the number system shall be generated by a number in the system ; the extension leads to the system of Real Numbers.

4. Real Numbers. The definition of the real number is as follows, it being understood that both of the classes L and U exist.

DEFINITION. If by any method *all* the rational numbers are separated into two classes L and U such that every number in L is less than every number in U, the section so determined is said to define a **real number**.

The symbol (L, U) may be used to denote either the real number or the section ; this double use of the symbol causes no confusion.

If there is a greatest number in L or a least in U then (L, U) corresponds to the *rational* number which generates the section and is called a *real rational number*. If L has no greatest number and U no least then (L, U) does not correspond to any rational number ; in this case (L, U) is called a *real irrational number*.

In regard to the terminology the following remarks of Hobson may be quoted (*Functions of a Real Variable*, 1st Ed., p. 29) :

"The rational numbers are frequently regarded as identical with the real numbers to which they correspond, and are denoted by the same symbols. In the development of Analysis, this identity leads to no difficulties ; but in the fundamental theory of the aggregate of real numbers, a conceptual distinction between rational numbers and the real numbers to which they correspond must be made, in order to obviate logical difficulties, and especially with a view to coordinating Cantor's theory with that of Dedekind. Those real numbers which do not correspond to rational numbers are called irrational numbers ; and those real numbers which correspond to rational numbers are usually spoken of as themselves rational numbers."

5. Properties of the Real Number. To save tedious repetitions, let the lower classes of sections be denoted by capital letters $A, B, C, \ldots$, and the corresponding upper classes by the

corresponding accented letters A', B', C', ... A typical (rational) number of A, B, C ... will be denoted by a, b, c, ... and a typical number of A', B', C', ... by a', b', c', ... When (A, A') is rational the generating number may be considered to belong either to A or to A', and the rational number that generates the section will be indicated by the suffix 0 attached to the typical letter. We adopt the convention that the generating number is to be the greatest number in the lower class; thus if (A, A'), (B, B'), (C, C') are rational real numbers the generating numbers of the sections are a_0, b_0, c_0 respectively. (See *Note*, below.)

The real numbers (A, A'), (B, B'), (C, C') ... whether rational or irrational will often be denoted by the corresponding Greek letters α, β, γ, ...

As yet the real number is little more than a name, and properties will now be assigned to it by means of definitions, the properties of rational numbers being assumed.

Positive Number. Negative Number. Zero.

The number (A, A') or α is defined to be positive if A contains a positive number, negative if A' contains a negative number. If all the numbers in A are negative and all those in A' positive, the section defines the number zero.

It may be noted that if A contains one positive number, k say, it contains an unlimited number of positive numbers; for $k/2$, $k/3$, ... are positive and being less than k are all in A.

Equality.

The number α or (A, A') is defined to be equal to the number β or (B, B'), and β to be equal to α, if the numbers in A are the same as those in B or—what amounts to the same thing—if the numbers in A' are the same as those in B'. In symbols: $\alpha=\beta$, $\beta=\alpha$.

Note. When the real number is rational the number that generates the section has been chosen to be the greatest number in the lower class, but this, of course, is merely a convention. If a_0 is the greatest number in A and b_0' the least number in B' the sections (A, A') and (B, B') will give the same real rational number if a_0 and b_0' are equal; the numbers defined by the sections must therefore be equal even though the class A

contains one number that is not in B (and B' one that is not in A'). To justify the convention it will be sufficient to show that if there is *only one number* in A, say a_1, that is not in B, then a_1 is at once the greatest number in A and the least in B' because when this is the case the sections (A, A') and (B, B') define the same real number.

Now since a_1 is not in B it must be in B'; let a_1 be denoted by b_1' when considered as a number in B'. But a_1 is the only number of A that lies in B' and therefore every other number in A lies in B and so is less than b_1', that is, less than a_1. Hence a_1 is the greatest number in A.

Again, every number less than b_1' is less than a_1 and, as has been seen, lies in B so that b_1' is the least number in B'. Since $a_1 = b_1'$ the sections (A, A') and (B, B') are generated by the same number and therefore define the same real number.

Thus, when the class B' contains a least number it may be transferred to the class B, which will then have a greatest number; this transference is always supposed to be made.

If there were *two* numbers in A that were not in B then there would be (§ 3, (2)) an unlimited number of such numbers and the sections (A, A') and (B, B') would be quite different.

Inequality.

If α and β are not equal α is defined to be greater than β and β less than α when A contains all the numbers in B *and more* (see above *Note*). In symbols: $\alpha > \beta$, $\beta < \alpha$.

It may be observed that if α and β are real rational numbers $\alpha > \beta$ or $\alpha < \beta$ according as $a_0 > b_0$ or $a_0 < b_0$, where a_0 and b_0 are the rational numbers that generate the sections (A, A') and (B, B'). (Compare Ex. 2 below.)

Further, it follows readily from the definitions that if $\alpha > \beta$ and $\beta > \gamma$ then $\alpha > \gamma$. (Compare § 3, (1).)

Next suppose $\alpha > \beta$. Let a be any rational number that is not in B and is therefore in B'. By convention, B' contains no least number and therefore there is an unlimited number of numbers in B' less than a; all these numbers are in A, so that, if r is any one of them and ϱ the corresponding real number, $\alpha > \varrho > \beta$. Hence, *between any two different real numbers lies an unlimited number of real rational numbers*. (Compare § 3, (2).)

Theorem. *If d_0 is any given arbitrarily small positive rational number it is always possible to find rational numbers, x and x', in the lower and upper classes respectively of a section (A, A') such that $x' - x < d_0$.*

Consider the arithmetical progression

$$a,\ a+d,\ a+2d,\ \ldots,\ a+nd \quad\ldots\ldots\ldots\ldots\ldots\ldots\text{(i)}$$

where a is any number in A and d is a positive rational number less than d_0. It is possible * to take n so large that nd shall be greater than $a'-a$, where a' is any number in A', and therefore so that $a+nd$ shall be greater than a' and therefore be in A'. Let $a+pd$ be the last number of the progression (i) that is in A; then $a+(p+1)d$ is in A' so that if $x=a+pd$ and $x'=a+(p+1)d$ we have $x'-x=d<d_0$, and the numbers x and x' satisfy the conditions of the theorem.

Ex. 1. Show that $\alpha>0$ if α is positive and $0>\alpha$ if α is negative.

Ex. 2. If (A, A') is the real number determined by assigning (as in § 3) to the upper class A' every rational number whose square is greater than 2 and to the lower class A all the other rational numbers, explain the meaning of the inequalities

$$1{\cdot}41<(A, A')<1{\cdot}42.$$

Here 1·41 and 1·42 are *real rational* numbers, determined by sections whose generating numbers are the *rational* numbers 1·41 and 1·42 respectively. The class A contains the rational numbers 1·414, 1·4142, ... and therefore the real number 1·41 is less than the real number (A, A'). Similarly the real number (A, A') is seen to be less than the real number 1·42.

A real number can only be compared, as respects the properties of "greater" and "less," with other real numbers. It may, however, be observed that the real rational number corresponds to the rational number in such a way that in inequalities like that given above we may simply take the rational numbers in the respective classes and *re-name* them, calling them real rational numbers, and then the comparison becomes valid. The farther we proceed the more evident it will become that the real rational number has no properties that are distinct from those of the corresponding rational number (compare the quotation from Hobson in § 4). See the *Note on Terminology* in § 7.

6. Laws of Operation. There is no question of "proving the laws"; the problem is to *prescribe* laws of operation that will be consistent with those that hold for rational numbers. The notation of § 5 is retained.

* The assumption that, if a and b are any two positive numbers and a less than b, an integer n can always be found such that na is greater than b, is usually called the *Axiom of Archimedes* (sometimes, perhaps more appropriately, the *Axiom of Eudoxus*. See Heath's *Euclid* (2nd Ed.), vol. iii. pp. 15, 16.).

Addition. Let a, a' and b, b' be typical numbers in the classes A, A' and B, B' respectively, and let $a+b=c$, $a'+b'=c'$; then $c<c'$. Now form the classes C and C' by assigning to C and C' all numbers of the type c and c' respectively.

If α and β are both rational a_0 and b_0 are the greatest numbers in A and B respectively; hence c_0, where $c_0=a_0+b_0$, is the greatest number in C so that (C, C') is a section generated by the rational number c_0. But c_0 is the sum of a_0 and b_0 and therefore, if the real number (C, C') is defined to be the sum of the real numbers (A, A') and (B, B'), the law of addition will be consistent with the corresponding law for rational numbers. If then the classes C and C' define one, and only one, number in all cases, the number (C, C') or γ will be defined to be the sum of (A, A') or α and (B, B') or β. In symbols

$$(A, A')+(B, B')=(C, C') \text{ or } \alpha+\beta=\gamma$$

It will now be proved that the classes C and C' always define one, and only one, real number. In all cases $c<c'$ and there cannot be more than one rational number that does not occur either in C or in C'. For if there were two, say x and y where $x<y$, then $c'-c$ could not be less than $y-x$. But, by the theorem of § 5, the numbers a, a' and b, b' can be chosen in their respective classes so that $a'-a<\frac{1}{2}d$, $b'-b<\frac{1}{2}d$, and therefore so that $c'-c$, which is equal to $(a'-a)+(b'-b)$, shall be less than d where d is arbitrarily small. Now d, and therefore $c'-c$, may be taken to be less than $y-x$; hence there cannot be more than one rational number that does not occur either in C or in C'. If there is none, (C, C') is a definite irrational number; if there is one, assign it to C and then (C, C') is a definite real rational number.

Subtraction. If $-A$ denotes the class of numbers obtained by changing the sign of every number in the class A, the number $(-A', -A)$ is defined to be the negative of (A, A'); that is,

$$(-A', -A)=-(A, A').$$

The operation of subtracting β from α is defined by the equation

$$\alpha-\beta=\alpha+(-\beta),$$

and is therefore reducible to addition.

Absolute Value of a Number. The absolute or numerical value of α is α if α is positive, $-\alpha$ if α is negative and zero if α is zero. The absolute value is denoted by $|\alpha|$.

Ex. 1. Prove that $\alpha - \alpha = 0$.

Ex. 2. Prove that $\beta - \alpha = -(\alpha - \beta)$.

Ex. 3. Prove that $|\alpha \pm \beta| \leqq |\alpha| + |\beta|$ but $\geqq \big| |\alpha| - |\beta| \big|$, and distinguish the cases in which the sign $=$ must be taken.

Multiplication. Suppose first that α and β are both positive and take (A, A') and (B, B') to be sections of the *positive* rational numbers; the conditions that in general determine a number will obviously, when all the positive rational numbers alone are taken, determine a positive number.

Let $ab = c$ and $a'b' = c'$; then $c < c'$. Form the classes C and C', as in defining addition; it is easy to prove as before that these classes determine one and only one real number. When α and β are rational a_0 and b_0 are the greatest numbers in A and B respectively and $a_0 b_0 = c_0$ so that the definition of multiplication, when α and β are positive, will be

$$(A, A') \times (B, B') = (C, C') \text{ or } \alpha\beta = \gamma.$$

To extend the definition to negative numbers assume the "rule of signs" *as part of the definition* so that we have, α and β being positive,

$$(-\alpha) \times \beta = -(\alpha \times \beta) = \alpha \times (-\beta); \; (-\alpha) \times (-\beta) = +(\alpha \times \beta).$$

Division. This operation is reduced to multiplication by first defining the reciprocal $1/\alpha$ of the positive number α.

Let α be determined by the classes A and A' of the positive rational numbers (zero excluded) and let $1/A$ and $1/A'$ denote the classes which contain the reciprocals of all the numbers in A and A' respectively. It is easy to prove that

$$(1/A', 1/A) \times (A, A') = 1,$$

and the number $(1/A', 1/A)$ is defined to be the reciprocal of (A, A'), that is,

$$(1/A', 1/A) = \frac{1}{\alpha} \text{ when } (A, A') = \alpha > 0.$$

If α is negative, $\alpha = -\alpha'$ where α' is positive, $1/\alpha$ is defined to be $-1/\alpha'$, that is, $1/\alpha = -1/(-\alpha)$ when α is negative.

The division of β by α is now defined to be the multiplication of β by the reciprocal of α; in symbols:

$$\frac{\beta}{\alpha}=\beta\times\left(\frac{1}{\alpha}\right).$$

Note. Division by zero is expressly excluded in the above definition.

The fundamental laws of operation have now been stated. A full treatment would go on to show that the associative, commutative and distributive laws of operation persist when the laws are defined as above, but in this sketch there is no room for the discussion and reference may be made to Chrystal's *Algebra* or Hobson's *Theory of Functions*. We may take one example to indicate the method when the real number is defined as in § 4.

To prove that $\alpha+\beta=\beta+\alpha$ note that the typical number $a+b$ in the lower class that defines $\alpha+\beta$ is equal to the typical number $b+a$ in the lower class that defines $\beta+\alpha$, since the commutative law holds for rational numbers. The lower classes are therefore the same for $\beta+\alpha$ as for $\alpha+\beta$, and similarly the upper classes are also the same. Hence the two numbers $\alpha+\beta$ and $\beta+\alpha$ are equal.

The student should prove the following cases in the same way.

Ex. 4. $\alpha\beta=\beta\alpha$.

Ex. 5. $(\alpha\beta)\times\gamma=\alpha\beta\gamma=\alpha\times(\beta\gamma)$.

Ex. 6. $(\alpha+\beta)\times\gamma=\alpha\gamma+\beta\gamma=\gamma(\alpha+\beta)$.

Ex. 7. $\alpha\times 0=0\times\alpha=0,\ \alpha\times 1=1\times\alpha=\alpha$.

Ex. 8. $|\alpha\beta|=|\alpha|\times|\beta|$.

7. Sections of the Real Numbers. The relations (1) and (2) stated for rational numbers in § 3 are true also for real numbers as follows from the definitions and developments stated in § 5. Hence the system of real numbers, which will be called the system S to distinguish it from the system R of rational numbers, is like R an *ordered* system. Further, like R, the system S is dense since it contains R. On the other hand, S possesses a property that is absent from R; namely, while there are sections of R that are not generated by a number in R, every section of S is generated by a number in S. By a

section of S is meant a separation of *all* the numbers in S into two classes, a lower class L and an upper class U, such that (i) both classes exist, (ii) every number in S appears either in L or in U, and (iii) every number in L is less than every number in U. That every section of S is generated by a number in S may be proved in the following way.

Take any section (L, U) of S. Let L_0 and U_0 contain all the rational numbers in R that correspond to real rational numbers in L and U respectively; then (L_0, U_0) is clearly a section of the system R and therefore defines a real number, that is, a number in S. If α is this number it must belong either to the class L or to the class U since L and U together contain all the numbers in S. Suppose α to belong to U and let β be any other number in U. By § 5 there are real rational numbers between α and β and, since these numbers correspond to numbers in U_0, they are all greater than α so that β is greater than α. Hence α is the least number in U.

In the same way it may be shown that if α belongs to L it is the greatest number in L. Thus in every section (L, U) of S either L has a greatest number or U has a least; it is not possible that there should be *both* a greatest number in L *and* a least in U, because all numbers between them would escape classification.

Hence every section of S is generated by a number in S; this property marks the essential distinction between the systems R and S and gives the character of continuity to the system of real numbers (see § 9).

Note on Terminology. Up to this point the distinction between the real rational number and the rational number of the system R—which may be called for convenience the "ordinary" rational number—has been preserved. The adherence to the distinction in the further development would, however, occasion intolerable prolixity and therefore the real numbers that correspond to the ordinary rational numbers will be called rational numbers, unless there be some special reason for emphasizing the distinction.

Even in the use of the ordinary rational numbers, as remarked in § 2, there is this use of the same term to indicate numbers that are conceptually distinct. Thus the numbers 2 and 2/1

are both said to be equal to "two," but the fraction $2/1$ is taken to be equivalent to the integer 2 as a convention or definition; the natural number 2 is not a quotient and would not be considered as a quotient except for reasons based on the development of the theory of fractions.

Again, there is no means of distinguishing whether symbols such as $1, 2, 1/2, -7, \ldots$ represent ordinary or real rational numbers but the context in which they appear will usually enable one to decide. If irrational numbers are associated with them the symbols must be interpreted in the sense of real rational numbers (see § 5, Ex. 2); if no irrational number is associated with them it does not matter which meaning is taken. But, as remarked in connection with Ex. 2 of § 5, any change needed amounts in actual work to a simple re-naming of the numbers since the real rational number always corresponds to an ordinary rational number, and the order of magnitude is the same for both, so that no confusion can occur.

The distinction of terminology between ordinary and real rational numbers will therefore, as a rule, be dropped. Further, any letter may be used to represent a real number, whether rational or irrational; the special use of Greek letters as denoting real numbers will therefore not be maintained.

Ex. 1. If the symbol $\sqrt{2}$ denotes the number (A, A') defined in § 5, Ex. 2, prove that $(\sqrt{2})^2$, that is, $(\sqrt{2})(\sqrt{2})$ is equal to 2.

If a and a' are typical numbers in A and A' then the classes of which a^2 and a'^2 are typical numbers determine a number, b say; b is greater than every a^2, less than every a'^2 and is therefore equal to 2.

Ex. 2. If $\sqrt{3}$ denotes the number (A, A') when 3 takes the place of 2 in § 5, Ex. 2, prove that $(\sqrt{2})(\sqrt{3}) = \sqrt{6}$ where 6 takes the place of 2 in § 5, Ex. 2.

8. Decimal Representation. Let α be a real number generating a section (L, U) of the system S of real numbers and, for definiteness, suppose α to be positive.

If a_0 is the greatest integer in L, then a_0+1 will be a number in U; a_0 may be zero. Thus we may write:

$$a_0 \leqq \alpha < a_0 + 1.$$

Next, form the arithmetical progression, with difference $1/10$,

$$a_0,\ a_0+\tfrac{1}{10},\ a_0+\tfrac{2}{10},\ \ldots,\ a_0+\tfrac{9}{10},\ a_0+1.$$

Of these eleven numbers one, which may be called $a_0 + a_1/10$ where a_1 is one of the numbers 0, 1, ... , 9, is the greatest in L, and then $a_0 + (a_1 + 1)/10$ is the least in U. Hence

$$a_0 + a_1/10 \leqq \alpha < a_0 + (a_1 + 1)/10.$$

Similarly, forming an arithmetical progression with $a_0 + a_1/10$ as its first term and with $1/10^2$ as common difference, we see that

$$a_0 + a_1/10 + a_2/10^2 \leqq a < a_0 + a_1/10 + (a_2 + 1)/10^2$$

where a_2 is one of the numbers 0, 1, ... , 9.

Proceeding in this way we find, in general, that

$$a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots + \frac{a_n}{10^n} \leqq \alpha < a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots + \frac{a_n + 1}{10^n};$$

or, say, $$\varrho_n \leqq \alpha < \varrho_n + \frac{1}{10^n} = \varrho'_n, \qquad \text{.......................(1)}$$

where each of the numbers $a_1, a_2, \dots, a_n$ has one of the values 0, 1, ... , 9, and $\varrho'_n - \varrho_n = 1/10^n$.

There are two cases to be considered.

(i) For $n > p$ we may have $a_n = 0$. In this case α is a rational number

$$a = a_0 \,.\, a_1 a_2 \dots a_p$$

in the usual decimal notation.

(ii) The set of numbers $a_1, a_2, \dots$ may be unlimited, and α will, in the usual terminology, be represented by an infinite decimal

$$\alpha = a_0 \,.\, a_1 a_2 a_3 \dots \qquad \text{..................(2)}$$

If the decimal "repeats" or "circulates" α will be rational; otherwise α is irrational. (See Chrystal's *Algebra*, Part II, 2nd Ed., Chap. 25, § 33.)

Cor. In the same way it may be proved that, if b is any positive integer not less than 2

$$\sigma_n \leqq \alpha < \sigma_n + 1/b^n = \sigma'_n$$

where $$\sigma_n = c_0 + \frac{c_1}{b} + \frac{c_2}{b^2} + \dots + \frac{c_n}{b^n}$$

and c_0 is an integer (or zero) while each of the numbers $c_1, c_2, \dots, c_n$ has one of the values 0, 1, ... , $(b-1)$.

Approximations. The numbers ϱ_n and ϱ'_n may, if $\alpha > \varrho_n$, be called rational approximations, in defect and in excess

respectively, to the number α, the absolute error in each approximation being less than $1/10^n$.

Between ρ_n and α there is, if $\alpha > \rho_n$, an unlimited number of real numbers, as also between α and ρ_n'; if x is any real number between ρ_n and α, and x' any real number between α and ρ_n', then

$$x' - x < \rho_n' - \rho_n, \text{ that is, } x' - x < \frac{1}{10^n}.$$

Since n may be taken so large that $1/10^n$ is less than ε, where ε is any arbitrarily small positive number, the following theorem is proved:

THEOREM. *It is always possible to find real numbers x and x' in the lower and upper classes respectively that define the real number α so that $x' - x < \varepsilon$, where ε is any arbitrarily small positive number.*

9. Correspondence of Number and Point. Suppose that, as in Analytical Geometry, O is the origin and OV the positive unit segment on an axis $X'OX$.

Let the points O and V correspond to the numbers 0 and 1 respectively.

If x is a positive rational number, say $x = m/n$ where m and n are positive integers, take a point P on the same side of O as V such that the segment OP is m times the nth part of OV; let the point P correspond to x. If x is negative $(= -m/n)$, take P' on the opposite side of O from V so that the segment OP' is of the same length as OP; let the point P' correspond to x. In the usual language of Analytical Geometry x is the abscissa of P. In this way the rational numbers are put into correspondence with points on the axis $X'OX$; for convenience let points which correspond to rational numbers be called "rational points."

If x is an irrational number, suppose it to be determined by a section (L, U) of the rational numbers and let A and A' be typical rational points corresponding to the typical numbers a and a' in L and U respectively. Now form a section of the points on the axis $X'OX$ by assigning to the lower (or left) class L_1 all points that correspond to rational points A or that lie to the left of any point A, and to the upper (or right) class U_1

all points that correspond to rational points A' or that lie to the right of any point A'.

In the class L_1 there is no point that lies furthest to the right since the class L contains no greatest number, and, similarly, since the class U contains no least number, there is no point in the class U_1 that lies furthest to the left. We now assume Dedekind's Axiom (§ 2) that there is one (and only one) point P on the axis $X'OX$ that generates the section (L_1, U_1) and we make P to correspond to x, so that x is the abscissa of P.

The correspondence between numbers and points on a line, assumed in § 4 of the *Elementary Treatise*, is therefore proved in so far as "proof" is possible. It is perhaps better simply to say that the system S of the real numbers forms a *continuum* or a *continuous system* of numbers, because every section of S is generated by a number in S, and that the continuity of the straight line is represented by the correspondence between the numbers in S and the points of the line: to rational numbers correspond rational points and to irrational numbers correspond irrational points.

The following definitions of terms that constantly occur may be given here.

Continuous Variable. The number x is said to vary *continuously* as it changes from the value a to the value b if, as it increases from a to b when $a < b$ or as it decreases from a to b when $a > b$, it takes every real value between a and b.

Interval. The system of numbers x such that $a \leqq x \leqq b$ is said to form a *closed* interval (a, b); the system of numbers x such that $a < x < b$ is said to form an *open* interval (a, b). The interval (a, b) is said to be "open at b" if $a \leqq x < b$ and "open at a" if $a < x \leqq b$.

The number ξ is said to be "within the interval (a, b)" or "interior to the interval (a, b)" if there are numbers x', x'' such that $a < x' \leqq \xi \leqq x'' < b$.

10. Roots. Indices. An important theorem will now be proved.

THEOREM. *If a is any positive real number and n a positive integer the equation $x^n = a$ has one, and only one, positive root.*

That there cannot be more than one positive root follows from the fact that if x and y are positive

$$x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + \ldots + xy^{n-2} + y^{n-1}),$$

and the product on the right cannot be zero unless $x = y$, since all the terms x^{n-1}, $x^{n-2}y$, ... are positive. Thus x^n and y^n cannot both be equal to a if x and y are two different positive numbers.

To prove that there is one root form a section of the *positive* rational numbers. (In the terminology the distinction between real rational number and "ordinary" rational number is dropped; see *Note* in § 7.) To the lower class L assign the rational number c if $c^n \leqq a$, and to the upper class U assign the rational number d if $d^n > a$. All the rational numbers are therefore classified and every number in the lower class is less than every number in the upper. The section therefore defines a real positive number, b say.

That $b^n = a$ follows at once from the definition of multiplication. The typical numbers in the lower and upper classes that define the product b^n are c^n and d^n, and it is merely a repetition of the process in § 6 to show that the section determined by the classes of which c^n and d^n are typical numbers defines a real number. b^n is the one number which is less than every number d^n and greater than or equal to any number c^n so that b^n and a are the same number.

This positive number b is the unique nth root of a and is denoted by the symbol $\sqrt[n]{a}$.

Cor. 1. If n is even, there is a second root but it is negative, namely $-b$.

Cor. 2. If n is odd and a negative, say $a = -a'$ where a' is positive, there is one negative root, namely, $-\sqrt[n]{a'}$.

In textbooks of algebra it is shown that when a is a positive real number the root $\sqrt[n]{a}$ may be denoted by the symbol $a^{1/n}$, and this index notation is then extended so that the symbol a^x has a definite meaning when x is any rational number. The complete symbol a^x is called a *power*; a is the *base* and x the *index* or *exponent* of the power. It has to be specially noted that the base a and the power a^x are both positive so that a^x is single-valued. For example, $4^{\frac{1}{2}}$ means $+2$ and not -2

even though the square of -2 is 4. Although a root such as $\sqrt[3]{(-8)}$ is a real number the notation $(-8)^{\frac{1}{3}}$ will not be used to indicate the root until the theory of the complex number is considered. With the conventions stated the power a^x is well defined when x is any rational number.

When the Theorem of this article has been proved the various laws of operation with rational indices, as developed for example in Chrystal's *Algebra*, are readily established, and it will be assumed that the student is familiar with them.

11. Inequalities. Some inequalities are frequently needed at later stages, and it seems to be desirable to state them here for reference.* They are based on the identity, n being a positive integer,

$$x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + \dots + xy^{n-2} + y^{n-1}). \quad \dots\dots(A)$$

It follows at once from (A) that if x and y are positive and n a positive integer

$$x^n > = < y^n \text{ according as } x > = < y. \quad \dots\dots\dots\dots(1)$$

Next, the identity (A) shows that if $x > y > 0$,

$$n(x-y)x^{n-1} > x^n - y^n > n(x-y)y^{n-1}. \quad \dots\dots\dots\dots\dots(2)$$

The inequalities (2) are very important; when n is not a positive integer the set requires a double statement which takes the following form:

If $x > y > 0$ and m a rational number

$$m(x-y)x^{m-1} > x^m - y^m > m(x-y)y^{m-1}, \begin{cases} m > 1 \\ \text{or } m < 0, \end{cases} \quad \dots\dots(3)$$

$$m(x-y)x^{m-1} < x^m - y^m < m(x-y)y^{m-1}, \quad 0 < m < 1. \quad \dots\dots(4)$$

The proof of (3) and (4) is a little tedious. In (2) let $y=1$, $x>1$, and the first of the inequalities gives $x^n - 1 < n(x-1)x^{n-1}$ and this inequality may be put in the forms

$$(n-1)(x^n - 1) > n(x^{n-1} - 1), \quad (x^n - 1)/n > (x^{n-1} - 1)/(n-1).$$

In the second form put $n-1, n-2, \dots, p+1$ successively in place of n; it follows that, p being a positive integer,

$$(x^n - 1)/n > (x^p - 1)/p, \quad n > p \geqq 1. \quad \dots\dots\dots\dots\dots\dots(\text{i})$$

Now in (i) let $x^p = a > 1$, $n/p = m > 1$; this gives

$$a^m - 1 > m(a-1), \quad a > 1, \quad m > 1. \quad \dots\dots\dots\dots\dots\dots(\text{ii})$$

while if $x^n = a > 1$, $p/n = m < 1$, we find from (i) that

$$a^m - 1 < m(a-1), \quad a > 1, \quad 0 < m < 1. \quad \dots\dots\dots\dots\dots(\text{iii})$$

* On the subject of this article the student should consult Chrystal's *Algebra*, Part II (2nd Ed.), Chapter XXIV.

Next, in the second of the inequalities (2) let $x=1$ so that $0<y<1$. The inequality $1-y^n>n(1-y)y^{n-1}$ may be put in the form

$$(1-y^n)/n<(1-y^{n-1})/(n-1),$$

and therefore, n, p being positive integers,

$$(1-y^n)/n<(1-y^p)/p, \quad n>p\geqq 1. \qquad \text{.....................(iv)}$$

Now let $y^p=b<1$, $n/p=m>1$, and (iv) gives

$$1-b^m<m(1-b), \quad b<1, \quad m>1, \qquad \text{.....................(v)}$$

while if $y^n=b<1$, $p/n=m<1$, formula (iv) gives

$$1-b^m>m(1-b), \quad b<1, \quad 0<m<1. \qquad \text{.................(vi)}$$

Put x/y for a in (ii) and y/x for b in (v); these substitutions give the inequalities (3) for the case $m>1$ while the same substitutions in (iii) and (vi) give the inequalities (4).

Finally, if $m>0$ put a for x, 1 for y and $m+1$ for m in formula (3) which has been established for these values; then

$$a^{m+1}-1<(m+1)(a-1)a^m \text{ gives } a^{-m}-1>-m(a-1)$$

$$a^{m+1}-1>(m+1)(a-1) \text{ gives } a^{-m}-1<-m(a-1)a^{-m-1},$$

so that if x/y is now put for a the formula (3) for the case $m<0$ is established.

From (3) another formula may be deduced. Suppose $a>0$, $b>0$, $m=-\mu(\mu>0)$ and r a positive integer. In the *first* of the inequalities (3) let $x=a+rb$, $y=a+(r-1)b$, and in the *second* $x=a+(r+1)b$, $y=a+rb$; then

$$\frac{1}{[a+(r-1)b]^\mu}-\frac{1}{(a+rb)^\mu}>\frac{\mu b}{(a+rb)^{\mu+1}}$$

$$>\frac{1}{(a+rb)^\mu}-\frac{1}{[a+(r+1)b]^\mu}. \qquad \text{.........(5)}$$

If $a>b>0$ the formula (5) holds when $r=0$.

The following particular cases of the formulae (3) and (4) may be noted: If $a>1$,

$$a^m>1 \text{ or } a^m<1 \text{ according as } m>0 \text{ or } m<0. \qquad \text{........(6)}$$

Further, since $a^y-a^x=a^x(a^{y-x}-1)$, it follows that if $a>1$,

$$a^y>a^x \text{ or } a^y<a^x \text{ according as } y>x \text{ or } y<x. \qquad \text{.........(7)}$$

Hence when x increases, taking rational values alone, the function a^x steadily increases; or, as x varies through *rational* values from $-N$ to $+N$, where N is a large positive number, a^x steadily increases from the small number a^{-N} (or $1/a^N$) to the large number a^N.

Again, the following simple particular cases of (3) and (4) may be stated.

If $h>0$, $(1+h)^m > 1+mh$ when $m>1$; $(1+h)^m < 1+mh$ when $0<m<1$, ...(8)

and if $0<h<1$ and $0<mh<1$

$$(1-h)^m > 1-mh > 0 \text{ when } m>1; \ (1-h)^m < 1-mh \text{ when } 0<m<1. \qquad \text{(9)}$$

The following inequalities are important in connection with the exponential function.

In (3) let $x=1+1/n$, $y=1$, $x-y=1/n$, $m=n/p$, where n and p are positive rational numbers and $n>p$; then

$$\left(1+\frac{1}{n}\right)^{\frac{n}{p}} > 1+\frac{1}{p};$$

and $$\left(1+\frac{1}{n}\right)^n > \left(1+\frac{1}{p}\right)^p, \ n>p>0. \qquad \text{(10)}$$

Next, in (3) let $x=1$, $y=1-1/n$, $x-y=1/n$, where m, n, p are as before except that $p>1$; then

$$\frac{1}{p} > 1-\left(1-\frac{1}{n}\right)^{\frac{n}{p}} \text{ or } \left(1-\frac{1}{n}\right)^n > \left(1-\frac{1}{p}\right)^p$$

and therefore $$\left(1-\frac{1}{n}\right)^{-n} < \left(1-\frac{1}{p}\right)^{-p}, \ n>p>1. \qquad \text{(11)}$$

From (10) it follows that $(1+1/z)^z$ steadily increases as z increases through positive rational values, and from (11) that $(1-1/z)^{-z}$ steadily decreases as z increases through positive rational values.

Again, in (11) let $p=2$ and for n put $n+1$; then

$$\left(1-\frac{1}{n+1}\right)^{-(n+1)} < 4 \text{ if } n>1. \qquad \text{(11}a\text{)}$$

But $$\left(1-\frac{1}{n+1}\right)^{-(n+1)} = \left(1+\frac{1}{n}\right)^{n+1} > \left(1+\frac{1}{n}\right)^n;$$

so that $$\left(1+\frac{1}{n}\right)^n < 4 \text{ if } n>1, \qquad \text{(10}a\text{)}$$

and $$\left(1-\frac{1}{n+1}\right)^{-(n+1)} - \left(1+\frac{1}{n}\right)^n = \frac{1}{n}\left(1+\frac{1}{n}\right)^n < \frac{4}{n} \text{ if } n>1. \qquad \text{(12)}$$

EXERCISES I.

1. If d is a positive integer but not the square of an integer show that d is not the square of a rational fraction.

[The following solution is given by Dedekind (*Stetigkeit u.s.w.*, pp. 13, 14).

If possible suppose d to be the square of a rational fraction t/u in its lowest terms (all numbers positive); then $t^2 - du^2 = 0$. There is always an integer λ such that $\lambda u < t < (\lambda + 1)u$ and therefore $t - \lambda u (= u'$ say) is a positive integer less than u. Let $t' = du - \lambda t$; then t' is also a *positive* integer ($du^2 - \lambda tu = t^2 - \lambda tu = tu'$). Now

$$t'^2 - du'^2 = (\lambda^2 - d)(t^2 - du^2) = 0,$$

and therefore $d = (t'/u')^2$. But $u' < u$ so that t/u is not in its lowest terms; the hypothesis made is thus untenable.]

2. If d is defined as in Ex. 1 let $y = x(x^2 + 3d)/(3x^2 + d)$, where x is a positive rational number and show that

$$y - x = \frac{2x(d - x^2)}{3x^2 + d}, \quad y^2 - d = \frac{(x^2 - d)^3}{(3x^2 + d)^2}.$$

Deduce that the section of the (positive) rational numbers, determined by assigning to the upper class all rational numbers whose square is greater than d and to the lower class all the other rational numbers, is not generated by a rational number.

3. Show that the formula for the nth root of d, corresponding to that of Ex. 2 for the square root, is

$$y = \frac{x\{(n-1)x^n + (n+1)d\}}{(n+1)x^n + (n-1)d}.$$

Apply the formula to calculate approximations to $\sqrt[n]{d}$.

[Prove that $x^n - d$ has the same sign as $x - y$ and that the product $(x^n - d)(y^n - d)$ is positive.]

4. If $y = (\lambda x + d)/(x + \lambda)$ where x, λ, d are positive rational numbers (d not a perfect square), prove that

$$y - x = \frac{d - x^2}{x + \lambda}, \quad y^2 - d = \frac{(\lambda^2 - d)(x^2 - d)}{(x + \lambda)^2}.$$

Hence show that if a/b is a rational approximation to $\sqrt{d}$, and λ the integer next greater than $\sqrt{d}$, a better approximation is given by $(\lambda a + bd)/(a + \lambda b)$, and the two approximations are either both greater or both less than $\sqrt{d}$.

5. Let a and d be two positive integers, the second not being a perfect square, and let $(a + \sqrt{d})^n = A_n + B_n\sqrt{d}$ where A_n and B_n are positive integers; show that

$$A_{n+1} = aA_n + dB_n, \quad B_{n+1} = aB_n + A_n,$$
$$A_n^2 - dB_n^2 = (a^2 - d)^n.$$

By taking $A_1 = a$, $B_1 = 1$, show that the fractions A_n/B_n for $n = 2, 3, \ldots$

give approximations to $\sqrt{d}$ of increasing accuracy. The value A_3/B_3 is that given by Ex. 2. $(a=x,\ A_3/B_3=y.)$

6. Let a be a rational approximation to $\sqrt{d}$ and let

$$a_1=\tfrac{1}{2}\left(a+\frac{d}{a}\right),\quad a_2=\tfrac{1}{2}\left(a_1+\frac{d}{a_1}\right),\ \dots\ a_n=\tfrac{1}{2}\left(a_{n-1}+\frac{d}{a_{n-1}}\right);$$

prove that $a_1, a_2, \dots$ are approximations to $\sqrt{d}$ of increasing accuracy, the approximations being in excess. Show that

$$\frac{a_n-\sqrt{d}}{a_n+\sqrt{d}}=\left(\frac{a_1-\sqrt{d}}{a_1+\sqrt{d}}\right)^{2^{n-1}}.$$

7. Prove that, if c is an approximation to $\sqrt[n]{a}$ in excess, c_1 where

$$c_1=\frac{1}{n}\left[(n-1)c+\frac{a}{c^{n-1}}\right],$$

is a closer approximation, also in excess.

Show that a still closer approximation, also in excess, is c_2 where

$$c_2=c_1-\frac{n-1}{2c}(c-c_1)^2.$$

[Let $a=c^n-k$, where k is positive and small, $0<k<c^n$; then

$$\sqrt[n]{a}=\sqrt[n]{(c^n-k)}=c(1-k/c^n)^{\frac{1}{n}}.$$

Expand by the Binomial Theorem. To find c_1 reject all powers of k above the first and put (c^n-a) in place of k. To find c_2 reject all powers of k above the second, and so on for closer approximations.

In calculating $\sqrt[n]{a}$ by this method we may begin by taking c to be the integer next greater than $\sqrt[n]{a}$. As a rule c_2 gives a sufficiently close approximation for ordinary needs, but of course the process can be repeated.]

8. If $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ are two sets of n numbers which may be either positive or negative, prove that

$$(\Sigma a_r b_r)^2=(\Sigma a_r^2)(\Sigma b_s^2)-\Sigma(a_r b_s-a_s b_r)^2,$$

where r and s take the values $1, 2, \dots, n$.

Deduce that

$$(\Sigma a_r b_r)^2 \leqq (\Sigma a_r^2)(\Sigma b_s^2).$$

The equation is usually referred to as **Lagrange's Identity** and the inequality as **Schwarz's Inequality.**

9. From the inequality on p. 173 of the *Elementary Treatise*,

$$x^p y^q<\left(\frac{px+qy}{p+q}\right)^{p+q},\quad x>0,\ y>0,\ x\neq y$$

where p, q are any positive rational numbers, deduce the inequalities (10) and (11) of § 11.

[$x=1+1/p,\ y=1$ gives (10) and $x=1,\ y=1-1/q\ (q>1)$ gives (11).]

10. Prove from the inequality in Ex. 9 that, if p and q are any positive rational numbers ($q > 1$),

$$\left(1+\frac{1}{p}\right)^p < \left(1-\frac{1}{q}\right)$$

[Let $x = 1 + 1/p$ and $y = 1 - 1/q$. This result gives the additional information that every value of $(1 + 1/p)^p$ is less than every value of $(1 - 1/q)^{-q}$, q being greater than unity so that $1 - 1/q$ may be positive.]

CHAPTER II

SETS. SEQUENCES. LIMITING POINTS. LIMITS

12. Sets. Sequences. In § 9 it has been shown that there is such a correspondence between the real numbers and the points on a directed line or axis that to each number corresponds a point and to each point corresponds a number—in other words, the correspondence is "one-to-one." This correspondence frequently enables us to simplify theorems by using the language of geometry, and the words "number" and "point" are often used as interchangeable; care must, however, be taken against the surreptitious substitution of geometrical for arithmetical conceptions in any demonstration.

A part of the system of real numbers or of points on a line is often called a set or **aggregate** of numbers or of points, and the numbers or points are spoken of as the "elements" of the set or aggregate. The set is said to be *infinite* if the number of elements in it is not limited, and *finite* if that number is finite. As finite sets are of little importance for our purposes, the word "set" will be understood to mean "infinite set" unless distinctly specified to be a "finite set."

A sequence is a particular case of a set. If to the integers $1, 2, \ldots, n, \ldots$ there correspond definite numbers $a_1, a_2, \ldots, a_n, \ldots$, the set $a_1, a_2, \ldots, a_n, \ldots$ is called a **sequence**; the element a_0, corresponding to 0, is frequently taken as the first element of the sequence. Thus in a sequence the elements are arranged in a definite order while in a set the order of the elements is indifferent.

To indicate a set or aggregate one may use a letter, S say, and specify the nature of the elements: for example, "the set S of all rational numbers" or "the set S of all rational numbers in the interval (0, 1)."

The sequence $a_1, a_2, \ldots, a_n, \ldots$ may be, and indeed is usually, denoted by enclosing the general element a_n in a parenthesis, thus (a_n); it is to be gathered from the context or from an explicit statement whether the first element of the sequence is a_1 or a_0 or some other element.

A set of numbers is said to be **bounded above** if there is a number K greater than every number of the set, **bounded below** if there is a number k less than every number of the set and simply **bounded** if both K and k exist.

The special numbers called "the upper bound" and "the lower bound" of a set are defined in the next article.

13. The Upper and Lower Bounds. In the rest of our work there will be frequent use of "an arbitrarily small positive number"; the letter ε will be reserved to indicate such a number so that it will be freely used in this meaning without further explanation, though of course occasions will arise when it may seem proper to state the meaning explicitly. When the symbol is used in any other sense an explicit statement of the meaning will be given if that meaning is not clear from the context.

Note 1. It may be noted in passing that if a and b are *constants* such that $|a-b|<\varepsilon$ we must have $a=b$; the proof of this assertion is "obvious."

Two theorems of fundamental importance will now be proved.

THEOREM I. *If a set S is bounded above there is a number H which has the following properties:* (i) *no number of the set is greater than H, and* (ii) *at least one number of the set is greater than $H-\varepsilon$.*

This number H is called the **upper bound** of the set S.

Since the set S is bounded above there is a number K greater than every number of the set. Two cases are possible.

(1) The set may contain one number, a say, that is greater than all the other numbers of the set. In this case $a=H$, because no number in S is greater than a, while a itself is greater than $a-\varepsilon$.

(2) No number of the set is greater than all the rest. In this case a section of the real numbers may be formed as follows:

Let x be any real number. If there is a number in S which is equal to or greater than x assign x to the lower class L, but if x is greater than every number in S assign x to the upper class U. Obviously both classes exist, since K and every number greater than K belongs to U while the numbers in S belong to L; further, every number in L is less than every number in U, and every real number occurs either in L or in U. Therefore the section (L, U) determines a number and this number is H, as may be shown in the following way.

By the construction of the section every number greater than H is in U and is also greater than every number in S, so that no number in S is greater than H. On the other hand, there is always a number in L, and therefore in S, which is greater than $H-\varepsilon$, because every number in L is equal to or less than *some* number in S.

Note 2. When the set S has no greatest number there is not only one but an unlimited number of the numbers of S between $H-\varepsilon$ and H because, if there were only a finite number, one of them would be the greatest of the set—contrary to the hypothesis that S contains no greatest number.

In the same way the following theorem is proved.

THEOREM II. *If a set S is bounded below there is a number h which has the following properties*: (i) *no number of the set is less than h and* (ii) *at least one number of the set is less than $h+\varepsilon$.*

This number h is called the lower bound of the set S. If S contains no least number there is an unlimited number of the numbers in S between h and $h+\varepsilon$.

If the set S is bounded (that is, bounded both above and below) both numbers H and h exist.

If a set S is not bounded above there is a number in the set greater than K, no matter how large the positive number K may be; in this case (by a stretch of language) the set is said to have $+\infty$ as its upper bound. In the same way if a set is not bounded below there is a number in the set less than $-K$, no matter how large the positive number K may be, and in this case the set is said to have $-\infty$ as its lower bound.

14. Limits. Notation. In the *Elementary Treatise*, § 41, two definitions of a limit are given, and the distinction between "limit" and "value" is pointed out; there seems to be no necessity for repeating the definitions here, but the student would do well to read the more important articles in Chapters IV and V and the earlier pages of Chapter XVII of the *Treatise*, as a knowledge of the working rules of limits is now assumed.

We shall now, however, use the symbol $\rightarrow$ to indicate that a variable tends to a limit; the symbol is due to the late Dr. Leathem, and is a very valuable improvement in notation. Thus, for example, we now write

$$\underset{x\rightarrow 0}{\mathcal{L}} \frac{\sin x}{x} = 1 \quad \text{instead of} \quad \underset{x=0}{\mathcal{L}} \frac{\sin x}{x} = 1,$$

or

$$\lim_{x\rightarrow 0} \frac{\sin x}{x} = 1 \quad \text{instead of} \quad \lim_{x=0} \frac{\sin x}{x} = 1.$$

Again, when it is said that "the sequence (a_n) tends to the limit l," what is meant is that "the variable a_n tends to the limit l when n tends to infinity"; in symbols

$$a_n \rightarrow l \text{ when } n \rightarrow \infty, \text{ or, } \underset{n\rightarrow\infty}{\mathcal{L}} a_n = l.$$

The sequence is understood to be an infinite sequence so that the explicit statement that n tends to infinity is hardly necessary; compare the corresponding expressions for series.

Null Sequence. When the sequence (a_n) tends to zero the sequence is frequently called a **Null Sequence.**

Finally, when it is said that a variable tends to a limit it is always to be understood that the limit is a *finite number* unless it is expressly stated or clearly implied that the variable tends to $+\infty$ or to $-\infty$. See *E.T.** §§ 40, 41.

15. Monotonic Functions. If $f(x)$ is a single-valued function of x such that $f(x_2) \geqq f(x_1)$ when $x_2 > x_1$ the function $f(x)$ is said to be (*E.T.* p. 451) a monotonic increasing function of x, while if $f(x_2) \leqq f(x_1)$ when $x_2 > x_1$ the function $f(x)$ is said to be a monotonic decreasing function of x. When the sign $=$ is

* The letters "*E.T.*" indicate that the reference is to the *Elementary Treatise*.

excluded $f(x)$ may be called "a strictly increasing" or "a strictly decreasing" function of x. (It is merely a convention, but a convenient one, to use the words "increasing" and "decreasing" instead of the more accurate descriptions "not decreasing" and "not increasing" respectively.) A sequence (a_n) is an increasing sequence if $a_{n+1} \geqq a_n$, and a decreasing if $a_{n+1} \leqq a_n$.

The two following theorems, which are expressed in terms of sequences (a_n), apply to monotonic functions $f(x)$, it being understood that $f(x)$ is defined for an infinite set of values of x and that x tends to infinity; the reasoning is the same in both cases.

THEOREM I. *A monotonic, increasing sequence (a_n) tends to $+\infty$ if the sequence is not bounded, but if the sequence is bounded above, say $a_n < k$ for every value of n, the sequence tends to a limit l, and $l \leqq k$.*

If the sequence is not bounded above then, however large the positive number K may be, it is always possible to find an integer m such that $a_n > K$ if $n > m$ (or a value x_1 of x such that $f(x) > K$ if $x > x_1$). This is the condition that a_n tend to $+\infty$ when $n \to \infty$ (or that $f(x) \to +\infty$ when $x \to +\infty$).

Suppose now that (a_n) is bounded above, and let H be the upper bound of the sequence (or of the set of values of $f(x)$). By the properties of the upper bound (§ 13) $a_n \leqq H$ for every value of n ($f(x) \leqq H$ for every value of x) and, given ε, there is a value, n_0 say, such that $a_{n_0} > H - \varepsilon$ (a value x_0 such that $f(x_0) > H - \varepsilon$). Hence, since $a_n \geqq a_{n_0}$ if $n > n_0$ ($f(x) \geqq f(x_0)$ if $x > x_0$), we have

$$H - \varepsilon < a_n \leqq H \text{ if } n > n_0 \ (H - \varepsilon < f(x) \leqq H \text{ if } x > x_0),$$

and this is the condition that $a_n \to H$ (or that $f(x) \to H$). Further, H cannot exceed k so that $l = H \leqq k$.

Cor. If $f(x)$ is monotonic and increases as x tends to a *monotonically*, $f(x)$ either tends to ∞ or tends to a limit

Let $x = a + (1/y)$ or $a - (1/y)$ according as x tends to a through values greater than a or through values less than a; then $x \to a$ when $y \to \infty$.

In the same way the next theorem is proved.

THEOREM II. *A monotonic, decreasing sequence* (a_n) *tends to* $-\infty$ *if the sequence is not bounded but, if the sequence is bounded below, say* $a_n > k$ *for every value of* n, *the sequence tends to a limit* l, *and* $l \geq k$.

Cor. If $f(x)$ is monotonic and decreases as x tends to a *monotonically*, $f(x)$ either tends to $-\infty$ or tends to a limit.

Ex. 1. $a_n = \frac{1}{3} - \frac{1}{3} \cdot \frac{1}{10^n}$.

Here we may take $k = 1/3$ and obviously $a_n \to 1/3$. (There is no value of n for which $a_n = 1/3$.)

Ex. 2. $a_n = \dfrac{1}{n+1} + \dfrac{1}{n+2} + \dfrac{1}{n+3} + \ldots + \dfrac{1}{2n}$.

Here $a_n < \dfrac{1}{n+1} + \dfrac{1}{n+1} + \ldots$ to n terms or $a_n < \dfrac{n}{n+1} < 1$, so that we may take $k = 1$. Further,

$$a_{n+1} - a_n = \frac{1}{2n+1} + \frac{1}{2n+2} - \frac{1}{n+1} = \frac{1}{(2n+1)(2n+2)} > 0,$$

so that (a_n) is an increasing sequence. The sequence therefore has a limit which is a positive number not greater than 1.

Ex. 3. If p is a fixed (positive) integer and a_n equal to the sum

$$\frac{1}{n+1} + \frac{1}{n+2} + \frac{1}{n+3} + \ldots + \frac{1}{n+pn},$$

show that the sequence (a_n) tends to a limit.

Ex. 4. If $a_m = \left(1 + \dfrac{1}{m}\right)^m$ and $b_n = \left(1 - \dfrac{1}{n}\right)^{-n}$ show that the sequences (a_m) and (b_n) tend each to a limit and that the limit is the same for both.

From the inequalities (10) and (11) of § 11 the sequences (a_m) and (b_n) are respectively increasing and decreasing; by the inequality (10*a*), $a_m < 4$ so that (a_m) tends to a limit, s say. Further

$$b_n = a_{n-1}\left(1 + \frac{1}{n-1}\right),$$

so that b_n tends to the same limit as a_m. It is better, however, to apply the result of Exercises I, 10, from which it appears that $b_n > a_m$ so that the sequence (b_n) is bounded below and therefore tends to a limit. The inequality (12) of § 11 then shows the two limits to be the same; or use *E.T.* p. 95 (iii).

Ex. 5. Suppose $a > b > 0$ and let $a_1 = \frac{1}{2}(a+b)$ and $b_1 = \sqrt{(ab)}$. If a_n and b_n are determined by the equations

$$a_n = \tfrac{1}{2}(a_{n-1} + b_{n-1}), \quad b_n = \sqrt{(a_{n-1}b_{n-1})}, \quad n = 2, 3, 4, \ldots$$

show that (i) (a_n) is a decreasing sequence, (ii) (b_n) is an increasing

sequence, (iii) each sequence tends to a limit and (iv) the limit is the same for each.

Here $a_{n-1} - a_n = \frac{1}{2}(a_{n-1} - b_{n-1})$, $b_n - b_{n-1} = (\sqrt{a_{n-1}} - \sqrt{b_{n-1}})\sqrt{b_{n-1}}$ and the results (i) and (ii) are easily proved by induction.

Again $a_n - b_n = \frac{1}{2}(\sqrt{a_{n-1}} - \sqrt{b_{n-1}})^2 > 0$ so that $a_n > b_n$. Now (b_n) is an increasing sequence and therefore $a_n > b$; the sequence (a_n) therefore tends to a limit α. Similarly $b_n < a$ and the sequence (b_n) tends to a limit β. The equation $a_n = \frac{1}{2}(a_{n-1} + b_{n-1})$ shows that $\alpha = \beta$.

The common limit α is called by Gauss the **Arithmetico-geometric Mean** of the numbers a and b, and is frequently denoted by $M(a, b)$.

Ex. 6. If $a > b > 0$, $a_1 = \frac{1}{2}(a+b)$, $b_1 = 2ab/(a+b)$ and a_n, b_n are determined by the equations

$$a_n = \tfrac{1}{2}(a_{n-1} + b_{n-1}),\quad b_n = 2a_{n-1}b_{n-1}/(a_{n-1} + b_{n-1}),\quad n = 2, 3, 4, \ldots$$

show that the sequences (a_n) and (b_n) have a common limit—called the **Arithmetico-harmonic Mean** of a and b.

Show that this mean is equal to $\sqrt{(ab)}$.

16. Sequence of Intervals. A useful method of determining a number depends on the construction of two monotonic sequences, one (a_n) an increasing and the other (b_n) a decreasing sequence; when a_n and b_n are represented as points on a directed line, say the x-axis of Coordinate Geometry, the segment I_n of the line which represents the interval (a_n, b_n) may, subject to certain conditions, be made to contract as n tends to infinity so as to determine a point.

The sequence (a_n, b_n) of intervals defines a number when the following conditions are satisfied:

(i) (a_n) is an increasing and (b_n) a decreasing sequence of real numbers;

(ii) $a_n < b_n$ for every value of n;

(iii) given an arbitrarily small positive number ε, there is an integer m such that $b_n - a_n < \varepsilon$ if $n \geqq m$.

That these conditions define a number is clear; for by (i) and (ii) the sequences (a_n) and (b_n) have limits, a and b say, while $b - a = \mathop{\mathcal{L}}(b_n - a_n) = 0$, by (iii), so that $b = a$. This number a is the number defined by the sequence (a_n, b_n).

In the geometrical interpretation every point a_n is to the left of every point b_n; the point a is either inside each of the intervals I_n or else, after a certain stage, at one end of each interval I_n, and, as n tends to infinity, the interval I_n

contracts, the end-points a_n and b_n tending from opposite sides to a.

It may happen that $a_n = a$ if $n > p$, or again that $b_n = a$ if $n > q$; in these cases a is an end-point of each interval I_n when n is greater than p or q respectively.

The number a will be said to be *common* to each interval (a_n, b_n), it being understood that each interval is closed.

For purposes of reference this method of determining a number may, for want of a better name, be called **the method of the decreasing interval.**

17. Limiting Points or Points of Condensation. Let S be an infinite set of numbers, or of points corresponding to them on a directed line.

DEFINITION. If there is a point ξ such that the interval $(\xi - \varepsilon, \xi + \varepsilon)$ contains an infinite number of points of the set, ξ is called a **limiting point**, or a **point of condensation**, of the set. As a number, ξ is a *limiting number* of the set.

ξ may be but is not necessarily a point of the set as the following examples show.

Ex. 1. S consists of all the rational numbers x such that $0 \leqq x \leqq 1$.
Every irrational number in the interval $(0, 1)$ corresponds to a limiting point—or, as it may be stated, every irrational point in $(0, 1)$ is a limiting point; but S contains no irrational points.

Ex. 2. S consists of all the points in the open interval $(0, 1)$.
The points 0 and 1 are limiting points but do not belong to S.

Ex. 3. S consists of all the points in the closed interval $(0, 1)$.
Every point of the interval is a limiting point.

Ex. 4. S consists of all the positive or negative integers.
There are in this case no limiting numbers; it is, however, sometimes said that $+\infty$ and $-\infty$ are limiting numbers of this set.

THEOREM I. *Every infinite bounded set has at least one limiting point.* (**The Bolzano-Weierstrass Theorem.**)

Apply the method of the decreasing interval. Since the set is bounded there is a number a less than every number in the set and a number b greater than every number in the set; let all the numbers be represented on an axis.

Bisect the interval (a, b). In one or, it may be, in both of the half-intervals there will be an infinite number of points of

the set; if there be an infinite number in both select the half-interval *on the right*, that is, the interval $[\frac{1}{2}(a+b), b]$. Denote the ends of the half-interval selected by a_1 and b_1 where $a_1 < b_1$. If the right half-interval has been selected $a_1 = \frac{1}{2}(a+b)$ and $b_1 = b$, but if the left half-interval has been chosen $a_1 = a$ and $b_1 = \frac{1}{2}(a+b)$. In both cases $a \leqq a_1$, $b \geqq b_1$, $b_1 - a_1 = \frac{1}{2}(b-a)$.

Next bisect the interval (a_1, b_1) and proceed exactly as in the preceding case. One at least of the new half-intervals contains an infinite number of points of the set, but if both half-intervals contain an infinite number select the half-interval *on the right* and call it (a_2, b_2). We now have

$$a \leqq a_1 \leqq a_2, \quad b \geqq b_1 \geqq b_2; \quad b_2 - a_2 = \tfrac{1}{2}(b_1 - a_1) = \frac{1}{2^2}(b-a).$$

Proceeding in this way we find an increasing sequence (a_n) and a decreasing sequence (b_n) where $a_n < b_n$, $(b_n - a_n) = (b-a)/2^n$ and each interval (a_n, b_n) contains an infinite number of points of the set. Further, since $(b-a)/2^n$ may be made arbitrarily small the sequence of intervals determines a point ξ.

Hence within the interval $(\xi - \varepsilon, \xi + \varepsilon)$, where ε has the usual meaning, there lies an infinite number of points of the set, and therefore ξ is a limiting point of the set.

Thus the set has at least one limiting point; but there may be more because, each time an interval is bisected, it is possible that both halves may contain an infinite number of points of the set. When an interval occurs which contains an infinite number of points of the set, that interval, by the theorem just proved, has at least one limiting point of the set.

The point ξ determined by the construction first given is, if there be more limiting points than one, *that which lies furthest to the right*; ξ is the greatest of the limiting numbers of the set and will be denoted by G. The characteristic property of G, as appears from the method by which it is determined, is that the interval $(G - \varepsilon, G + \varepsilon)$ contains an infinite number of points of the set but that only a finite number (and there may be none) of the points of the set lies to the right of $G + \varepsilon$.

The number G is called *the greatest of the limiting numbers* of the set or, simply, the *greatest of the limits* of the set.

When there are more limiting points than one of the set it may be seen that, if in the construction the *left* half-interval is

always selected instead of the right half, there is a limiting point g such that the interval $(g-\varepsilon, g+\varepsilon)$ contains an infinite number of points of the set, but that only a finite number (and there may be none) of the points lies to the left of $g-\varepsilon$. The number g is the *least of the limits* of the set.

THEOREM II. *If the set is bounded and has no greatest number the upper bound H is a limiting number of the set and $H=G$, while if the set has no least number the lower bound h is a limiting number of the set, and $h=g$.*

If the set has no greatest number then the interval $(H-\varepsilon, H)$ contains an infinite number of elements of the set so that H is a limiting number. Again no number in the set is greater than H so that $H=G$. Similarly for the lower bound.

The greatest and the least of the limits are often called the *Maximum Limit* and the *Minimum Limit* respectively of the set. Other names are *limes superior* (or upper limit) and *limes inferior* (or lower limit).

When the set is denoted by S the following notations are used:

$$G=\overline{\lim}\ S \text{ or } G=\text{lim. sup. } S\ ;\ g=\underline{\lim}\ S \text{ or } g=\text{lim. inf. } S.$$

When S is a sequence (a_n) the notations are

$$G=\overline{\lim_{n\to\infty}}\ a_n=\overline{\lim}\ a_n\ ;\ \ g=\underline{\lim}_{n\to\infty}\ a_n=\underline{\lim}\ a_n,$$

the indication "$n\to\infty$" being omitted when no ambiguity arises.

The notation $\overline{\underline{\lim}}$ is sometimes used when it is a matter of indifference whether the maximum or the minimum limit is taken. For example, the inequalities

$$a<\overline{\underline{\lim}}\ a_n<b,$$

imply that $a<g$ and $G<b$.

Note on Sequences selected from Sets. If c is a limiting point of a set S that lies in an interval (a, c), where $a<c$, a monotonic, increasing sequence (x_n) can be selected from S such that (x_n) tends to c. This statement seems to be "obvious," but a formal proof may be given.

If x_1, where $x_1>a$, is any point of the set and $a_1=(x_1+c)/2$, infinitely many points of S lie in (a_1, c) and, if x_2 is any one of these points, $x_2>x_1$ and $c-x_2<(c-a)/2$. Again, if $a_2=(x_2+c)/2$, let x_3 be one of the infinitely many points of S in (a_2, c); then $x_3>x_2$ and $c-x_3<(c-a)/2^2$. Next, let $a_3=(x_3+c)/2$ and a point x_4 may be chosen from the points

of S in (a_3, c) such that $x_4 > x_3$ and $c - x_4 < (c-a)/2^3$. In this way a monotonic, increasing sequence (x_n) is obtained and $(x_n) \to c$, since $c - x_n < (c-a)/2^{n-1}$. Clearly the sequence may be chosen in infinitely many ways.

Similarly, if c is a limiting point of a set S' that lies in (c, b), where $c < b$, a monotonic, decreasing sequence (x'_n) can be selected (in infinitely many ways) from S' such that (x'_n) tends to c.

If c is a limiting point for both of the sets S and S' let $x''_{2n-1} = x_n$ and $x''_{2n} = x'_n$; then the sequence (x''_n) tends to c, so that from the set S'', consisting of the sets S and S' and lying in the interval (a, b), a sequence (x''_n)—of course, not monotonic—has been selected which tends to c, where $a < c < b$ and c is a limiting point of S''.

18. Limits of Indetermination. The maximum limit G and the minimum limit g of a bounded sequence (a_n) are sometimes called respectively *the upper limit of indetermination* and *the lower limit of indetermination* of the function a_n for n tending to ∞.

Every element a_n of the sequence (a_n) lies between H and h, the upper and lower bounds of the sequence but, when the question of a *limit* for a_n arises, the only values of a_n that are of importance are those corresponding to large values of n, and it is on these large values that the limit depends. Now, when ε is given, there is only a finite number of values a_n that are greater than $G + \varepsilon$ while there is an infinite number greater than $G - \varepsilon$; so also there is only a finite number of values of a_n less than $g - \varepsilon$ but an infinite number less than $g + \varepsilon$. By § 17, Theorem II, $G = H$ provided (a_n) has no greatest number, and $g = h$ provided (a_n) has no least.

Next let $f(x)$ be a bounded, single-valued function of x, defined for an infinite set of values (not necessarily *all* values) of x in an interval $(\xi - \varepsilon, \xi + \varepsilon)$, ξ being a limiting point of the set of values x. Since $f(x)$ is bounded the set of values of $f(x)$ is a bounded set and has therefore a maximum limit G and a minimum limit g.

Now let ε tend to zero. It is easy to prove, if it be not considered to be obvious, that as ε tends to zero G cannot increase and g cannot decrease, so that G and g are bounded monotonic functions of ε, say $G(\varepsilon)$ and $g(\varepsilon)$. It follows therefore from § 15 that $G(\varepsilon)$ and $g(\varepsilon)$ tend to limits, G_0 and g_0 say, when ε tends to zero.

These two numbers G_0 and g_0 are called respectively *the upper and the lower limit of indetermination of $f(x)$ for x tending to ξ.*

If the interval for which $f(x)$ is defined is $(\xi, \xi+\epsilon)$ then x can only tend to ξ through values greater than ξ or—in the usual notation—x tends to $\xi+0$. As before G and g tend to limits, G_0' and g_0' say, when $x \to \xi+0$; G_0' and g_0' are called the *right-hand* upper and lower limits of indetermination of $f(x)$ for x tending to $\xi+0$.

Similarly if the interval is $(\xi-\epsilon, \xi)$, so that x tends to ξ through values less than ξ $(x \to \xi - 0)$, there are *left-hand* upper and lower limits of indetermination of $f(x)$ for x tending to $\xi-0$; these may be denoted by G_0'' and g_0'' respectively.

In all these cases it must be remembered (i) that there is no sense in speaking of a limit for x tending to ξ unless ξ is a limiting point of the values of x for which the function is defined, and (ii) that the limits G_0 and g_0 depend on values of $f(x)$ for values of x such that $0 < |\xi - x| < \delta$ where δ is positive and arbitrarily small.

Again the difference $G-g$ is never negative; when $\epsilon \to 0$, then $G \to G_0$, $g \to g_0$, and therefore

$$0 \leqq G_0 - g_0 \leqq G - g.$$

The conditions that (a_n) should tend to a limit when $n \to \infty$ and that $f(x)$ should tend to a limit when $x \to \xi$ may be readily derived from the above statements about the limits of indetermination. The conditions in these cases are, however, of so fundamental a character that they will be considered in detail in §§ 19 and 21.

19. Existence of a Limit. Sequence. We shall now prove Cauchy's Test for the existence of a limit of the sequence (a_n) (*E.T.* p. 378, Th. III); the general case for the limit of a function of x is considered in § 21.

THEOREM. *The necessary and sufficient condition that the sequence (a_n) tend to a limit, l say, is that, given an arbitrarily small positive number ϵ, there shall be an integer m such that $|a_{n+p} - a_n| < \epsilon$ if $n \geqq m$, whatever value the integer p may take.*

(i) The condition is necessary. For if $a_n \to l$ there is, by the

definition of a limit, an integer m such that $|l-a_n|<\frac{1}{2}\varepsilon$ if $n \geqq m$. But

$$|a_{n+p}-a_n| = |(a_{n+p}-l)+(l-a_n)| \leqq |l-a_{n+p}| + |l-a_n|$$

and therefore

$$|a_{n+p}-a_n| < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon, \text{ that is, } |a_{n+p}-a_n| < \varepsilon \text{ if } n \geqq m,$$

so that the condition is necessary.

(ii) The condition is sufficient. For, if the condition is satisfied there is an integer m such that if $n \geqq m$

$$|a_{n+p}-a_n| < \varepsilon, \text{ or, } a_n - \varepsilon < a_{n+p} < a_n + \varepsilon, \; p = 1, 2, 3, \ldots,$$

and therefore the sequence (a_μ), where μ take the values, $m+1$, $m+2$, $m+3$, ..., is bounded and has a maximum limit G and a minimum limit g. In other words, if $n \geqq m$,

$$a_n - \varepsilon \leqq g \leqq G \leqq a_n + \varepsilon,$$

so that

$$0 \leqq G - g \leqq 2\varepsilon.$$

Hence, since ε is arbitrarily small, $G=g$, and there is only one limiting point in the sequence. The sequence (a_n) therefore tends to a limit; the limit is l, where $l=G=g$.

The student should prove that the other method, specified in the enunciation of Theorem III (*E.T.* p. 378), of stating the condition is equivalent to the above.

20. Examples. The following examples contain some interesting theorems in limits; others will be found in the Exercises at the end of the chapter. In some cases a knowledge of the chief theorems in the convergence of series is assumed.

Ex. 1. If a_n and b_n tend to zero when $n \to \infty$ and if further (b_n) is a monotonic decreasing sequence so that $b_n > b_{n+1} > 0$ (at least for sufficiently large values of n), then

$$\mathop{\mathcal{L}}_{n\to\infty} \frac{a_n}{b_n} = \mathop{\mathcal{L}}_{n\to\infty} \frac{a_n - a_{n+1}}{b_n - b_{n+1}},$$

provided that the second limit exists, whether that limit is a finite number l or infinite.

(i) Suppose that $(a_n - a_{n+1})/(b_n - b_{n+1})$ tends to a finite limit l. In this case there is an integer m such that if $n \geqq m$

$$\left| l - \frac{a_n - a_{n+1}}{b_n - b_{n+1}} \right| < \epsilon, \quad \text{or} \quad l - \epsilon < \frac{a_n - a_{n+1}}{b_n - b_{n+1}} < l + \epsilon,$$

or, since $b_n - b_{n+1}$ is positive, (m may be taken large enough to make each b_n positive)

$$(l-\epsilon)(b_n - b_{n+1}) < a_n - a_{n+1} < (l+\epsilon)(b_n - b_{n+1}). \quad \ldots\ldots\ldots\ldots(\alpha)$$

In (a) put $n+1, n+2, \ldots, n+p-1$ in turn for n and add; then for every integer p

$$(l-\epsilon)(b_n - b_{n+p}) < a_n - a_{n+p} < (l+\epsilon)(b_n - b_{n+p}), \quad n \geqq m.$$

Now let p tend to ∞; then $a_{n+p} \to 0$, $b_{n+p} \to 0$, and therefore

$$(l-\varepsilon)b_n \leqq a_n \leqq (l+\varepsilon)b_n, \quad n \geqq m,$$

or, dividing by the positive number b_n,

$$l-\varepsilon \leqq \frac{a_n}{b_n} \leqq l+\varepsilon, \text{ or } \left| l \quad \right| \leqq \varepsilon, \quad n \geqq m.$$

Hence $a_n/b_n \to l$ when $n \to \infty$.

(ii) Suppose that $(a_n - a_{n+1})/(b_n - b_{n+1})$ tends to ∞. In this case there is an integer m such that, given any positive number K, if $n \geqq m$,

$$(a_n - a_{n+1})/(b_n - b_{n+1}) > K, \text{ or, } a_n - a_{n+1} > K(b_n - b_{n+1}),$$

since $(b_n - b_{n+1})$ is positive. Proceeding as in case (i) we find that

$$a_n - a_{n+p} > K(b_n - b_{n+p}), \ a_n \geqq Kb_n, \ n \geqq m$$

and therefore $a_n/b_n \geqq K$ if $n \geqq m$. Hence $a_n/b_n \to \infty$ when $n \to \infty$ since K is any positive number.

Ex. 2. If $b_{n+1} > b_n$ and if $b_n \to \infty$ when $n \to \infty$, then

$$\underset{n\to\infty}{\mathcal{L}} \frac{a_n}{b_n} = \mathcal{L} \frac{a_{n+1} - a_n}{b_{n+1} - b_n},$$

provided that the second limit exists, whether that limit is a finite number l or infinite.

(i) Suppose the limit of $(a_{n+1} - a_n)/(b_{n+1} - b_n)$ to be the finite number l. As in the proof of Ex. 1 we find, since $b_{n+1} > b_n$, that

$$(l-\varepsilon)(b_{n+p} - b_n) < a_{n+p} - a_n < (l+\varepsilon)(b_{n+p} - b_n) \text{ if } n \geqq m,$$

or, writing a_n, b_n in place of a_{n+p}, b_{n+p}, and a_m, b_m in place of a_n, b_n,

$$(l-\varepsilon)(b_n - b_m) < a_n - a_m < (l+\varepsilon)(b_n - b_m) \text{ if } n > m.$$

Now divide by the positive number b_n and then add a_m/b_n to each member of these inequalities; this gives, if $n > m$,

$$\left(l-\varepsilon\right)\left(1 - \frac{b_m}{b_n}\right) + \frac{a_m}{b_n} < \frac{a_n}{b_n} < \left(l+\varepsilon\right)\left(1 - \frac{b_m}{b_n}\right) + \frac{a_m}{b_n}.$$

Let the numbers a_m, b_m be kept fixed and let n, which is greater than m, tend to infinity; since b_n tends to infinity with n it is therefore possible to choose N so that

$$l - 2\varepsilon < a_n/b_n < l + 2\varepsilon \text{ if } n \geqq N,$$

and therefore a_n/b_n tends to l when $n \to \infty$.

(ii) Suppose that $(a_{n+1} - a_n)/(b_{n+1} - b_n)$ tends to ∞; then, as before, we can find m so that, whatever positive number K may be, if $n > m$,

$$a_n - a_m > K(b_n - b_m) \text{ or } \frac{a_n}{b_n} > K\left(1 - \frac{b_m}{b_n}\right) + \frac{a_m}{b_n}.$$

Keeping a_m and b_m fixed, we can choose N so that a_n/b_n shall be

greater than $K-\varepsilon$ if $n>N$. Hence $a_n/b_n \to \infty$ when $n \to \infty$ since K is any positive number.

Compare *E.T.* p. 420, paragraphs I and II, with the theorems of Ex. 1 and Ex. 2.

Ex. 3. Prove the following theorems, usually called **Cauchy's First and Second Theorems.**

First Theorem. If $a_1+a_2+\ldots+a_n=s_n$, and if $a_n \to l$ when $n \to \infty$ so does s_n/n, where s_n/n is the Arithmetic Mean of the numbers $a_1, a_2, \ldots, a_n$.

Second Theorem. If $a_1, a_2 \ldots a_n$ are all positive and if a_n/a_{n-1} tends to a limit, r say, when $n \to \infty$, then $\sqrt[n]{a_n}$ also tends to r when $n \to \infty$.

The First Theorem is a particular case of Ex. 2; in that example let $a_n=s_n$, $b_n=n$, and we get the First Theorem.

Next write a_n in the form

$$\overline{1} \qquad a_1$$

therefore $$\log(\sqrt[n]{a_n})=\frac{\log(a_1/1)+\log(a_2/a_1)+\ldots+\log(a_n/a_{n-1})}{n}.$$

Now $a_n/a_{n-1} \to r$ and therefore, since $\log x$ is a continuous function of x, $\log(a_n/a_{n-1}) \to \log r$. By the First Theorem the fraction to which $\log(\sqrt[n]{a_n})$ is equal tends to $\log r$ so that $\sqrt[n]{a_n}$ tends to r.

It should be observed that s_n/n *may tend to a limit though* s_n *does not.* For example, let $a_n=(-1)^{n-1}$; s_n does not tend to a limit but the mean s_n/n tends to zero.

Ex. 4. If the sequences (a_n) and (b_n) tend to a and b respectively then the sequence (c_n), where

$$c_n=\frac{1}{n}(a_1b_n+a_2b_{n-1}+a_3b_{n-2}+\ldots+a_nb_1)$$

tends to ab.

Let $a_r=a+d_r$; then

$$c_n=a\,.\,\frac{1}{n}(b_n+b_{n-1}+\ldots+b_1)+\frac{1}{n}(d_1b_n+d_2b_{n-1}+\ldots+d_nb_1).$$

By Cauchy's First Theorem $(b_n+b_{n-1}+\ldots+b_1)/n$ tends to b. Next each of the numbers b_r is finite, say $|b_r|<B$, while $d_n \to 0$; therefore

$$(d_1b_n+d_2b_{n-1}+\ldots+d_nb_1) \quad < B\,\frac{|d_1|+|d_2|+\ldots+|d_n|}{},$$

and, again by the First Theorem, this expression tends to zero so that (c_n) tends to ab.

21. Existence of a Limit. Function of x. We now take the general case of Cauchy's Test for the Existence of a Limit; the theorem of § 19 is limited to the case of a sequence.

Let $f(x)$ be a bounded, single-valued function of x, defined for an infinite set of values of x in an interval (a, b); it is to be

understood that in the following discussion the values assigned to x are those values in (a, b)—*admissible* values they may be called—for which $f(x)$ is defined. Since these admissible values form an infinite set there is at least one limiting point ξ of the set. The theorem to be proved is as follows :

THEOREM. *The necessary and sufficient condition that $f(x)$ tend to a limit, l say, when x tends to ξ is that, given an arbitrarily small positive number ε, there shall be a positive number η such that $|f(x')-f(x'')|<\varepsilon$, when x' and x'' are any two values of x, such that $0<|x'-\xi|<\eta$ and $0<|x''-\xi|<\eta$.*

It should be noted that ξ itself is not an admissible value of x.

(i) The condition is necessary. For if $f(x)\to l$ when $x\to\xi$ it is possible to choose η so that $|l-f(x')|<\frac{1}{2}\varepsilon$ and $|l-f(x'')|<\frac{1}{2}\varepsilon$ when $0<|x'-\xi|<\eta$ and $0<|x''-\xi|<\eta$. Now

$$|f(x')-f(x'')|=|\{f(x')-l\}+\{l-f(x'')\}|\leqq|l-f(x')|+|l-f(x'')|$$

and therefore

$$|f(x')-f(x'')|<\varepsilon \text{ if } 0<|x'-\xi|<\eta \text{ and } 0<|x''-\xi|<\eta,$$

so that the condition is necessary.

(ii) The condition is sufficient. For, if the condition is satisfied there is a number η such that

$$|f(x')-f(x'')|<\varepsilon, \text{ or, } f(x'')-\varepsilon<f(x')<f(x'')+\varepsilon$$

when $$0<|\xi-x'|<\eta \text{ and } 0<|\xi-x''|<\eta.$$

Thus the set of values of $f(x)$ obtained by assigning to x all the admissible values of x in $(\xi-\eta,\ \xi+\eta)$ is a bounded infinite set and has therefore a maximum limit G and a minimum limit g. Hence

$$f(x'')-\varepsilon\leqq g\leqq G\leqq f(x'')+\varepsilon$$

so that $$0\leqq G-g\leqq 2\varepsilon.$$

But, as η tends to zero, G and g (see § 18) tend respectively to G_0 and g_0, where $G_0\leqq G$ and $g_0\geqq g$; therefore

$$0\leqq G_0-g_0\leqq 2\varepsilon.$$

As ε is arbitrarily small it follows that $G_0=g_0$ and therefore, when $x\to\xi$, the set of admissible values of $f(x)$ has only one limiting number G_0 or g_0 and $l=G_0=g_0$.

If ξ is a limiting point for values of x greater than ξ, similar reasoning shows that $f(x)$ will tend to a limit when $x\to\xi+0$

if, and only if, there is a number η such that $|f(x')-f(x'')|<\varepsilon$ when $0<|\xi-x'|<\eta$ and $0<|\xi-x''|<\eta$ $(x'>\xi,\ x''>\xi)$. If ξ is a limiting point for values of x less than ξ then $f(x)$ will tend to a limit when $x\to\xi-0$ if, and only if, there is a number η such that $|f(x')-f(x'')|<\varepsilon$ when $0<|\xi-x'|<\eta$, $0<|\xi-x''|<\eta$ $(x'<\xi,\ x''<\xi)$.

Again to find the condition that $f(x)$ should tend to a limit when $x\to+\infty$ let $x=1/y$ and $f(x)=F(y)$. When $x\to+\infty$ the new variable y tends to zero and the condition

$$|F(y')-F(y'')|<\varepsilon \text{ if } 0<y'<\eta \text{ and } 0<y''<\eta$$

becomes

$$|f(x')-f(x'')|<\varepsilon \text{ if } x'>N \text{ and } x''>N,$$

where $N=1/\eta$, and therefore N is an arbitrarily large positive number.

Similarly if $x\to-\infty$ the condition for a limit of $f(x)$ is

$$|f(x')-f(x'')|<\varepsilon \text{ if } x'<-N \text{ and } x''<-N,$$

where N is an arbitrarily large positive number.

Continuity of $f(x)$. If $f(x)$ is defined for *all* values of x in the interval $(\xi-\eta,\ \xi+\eta)$ the condition that $f(x)$ should be continuous at ξ is (*E.T.* p. 87) that $f(x)$ should tend to $f(\xi)$ when x tends to ξ; in other words that the *limit* of $f(x)$ for x *tending* to ξ should be equal to the *value* of $f(x)$ for x *equal to* ξ. This condition may now be stated in another form.

Let (x_n) be any sequence that tends to ξ; if the sequence

$$f(x_1),\ f(x_2),\ f(x_3),\ \dots$$

tends to $f(\xi)$ whatever particular sequence (x_n) be chosen, then $f(x)$ will be continuous at ξ.

That the two forms of the condition are equivalent may be proved as follows.

(*a*) If $(x_n)\to\xi$, then m may be chosen so that $|\xi-x_n|<\eta$ when $n\geqq m$; hence if $|f(\xi)-f(x)|<\varepsilon$ when $|\xi-x|<\eta$, we shall have $|f(\xi)-f(x_n)|<\varepsilon$ when $n\geqq m$, so that $f(x_n)\to f(\xi)$ when $(x_n)\to\xi$. Thus, when the first form of the condition is satisfied, so is the second.

(*b*) When the second form is satisfied, so is the first. For, if $f(x)$ does not tend to $f(\xi)$ when x tends to ξ, it will be possible to choose ϵ so that $|f(\xi)-f(x)|\geqq\varepsilon$ for *infinitely many values of* x (for a set S, say) in the interval $(\xi-\eta,\ \xi+\eta)$; because, if there were only a finite number of such values of x, one of them (x', say) would differ from ξ by less than

any of the others and then $|f(\xi)-f(x)|$ would be less than ε when $|\xi-x)|<\eta'$, where $\eta'=|\xi-x'|$, so that $f(x)$ would tend to $f(\xi)$ when x tended to ξ.

Now from the set S in the interval $(\xi-\eta, \xi+\eta)$ a sequence (x_n) can be selected which tends to ξ (§ 17, *Note*); hence $|f(\xi)-f(x_n)|\geqq\varepsilon$ for every value of n, so that $f(x_n)$ does not tend to $f(\xi)$ when (x_n) tends to ξ. This conclusion contradicts the condition that *every* sequence $f(x_n)$ tends to $f(\xi)$ when $(x_n)\to\xi$; it is therefore possible, when the second condition is satisfied, to choose η so that $|f(\xi)-f(x)|<\varepsilon$ when $|\xi-x|<\eta$. Thus, when the second form is satisfied, so is the first.

The student may, as an exercise, deduce the conditions in § 21 from those in § 19.

22. Exponential Functions. The exponential function a^x is, up to this stage, defined for rational values of the variable x alone. The definition will now be extended to irrational values of x by showing that if (x_n) is any sequence of rational numbers that has x as its limit the sequence (a^{x_n}) also has a limit and this limit is defined to be the value of a^x. The discussion is rather long but it is not difficult; the inequalities of § 11 are required. It has to be remembered that the base a is *positive*.

(1) $a^x\to1$ when x tends to zero through rational values.

Suppose $a>1$ and let $a^{1/n}=1+a_n$ where a_n is positive and n a positive integer; then (§ 11, (8))

$$a=(1+a_n)^n>1+n\,a_n,\ a_n<(a-1)/n$$

and therefore $a_n\to0$ and $a^{1/n}\to1$ when $n\to\infty$.

Further, $a^{-1/n}=1/a^{1/n}$ and therefore $a^{-1/n}\to1$ when $n\to\infty$.

Next, let (x_n) be any sequence of rational numbers that tends to zero and let N be an arbitrarily large positive integer. It is possible to choose m so that x_n will lie between $-1/N$ and $1/N$ if $n>m$, and therefore also so that a^{x_n} lies between $a^{-1/N}$ and $a^{1/N}$ if $n>m$. But, given ε as usual, N may by the preceding part of the proof be taken so large that both $a^{-1/N}$ and $a^{1/N}$ lie within the interval $(1-\varepsilon, 1+\varepsilon)$; so therefore does a^{x_n} if $n>m$, and therefore $a^{x_n}\to1$ when $x_n\to0$.

Suppose next that $a<1$. Then $b=1/a>1$ and $a^{x_n}=1/b^{x_n}$ so that $a^{x_n}\to1$ since $b^{x_n}\to1$ if $n\to\infty$.

(2) Next let (x_n) be a sequence of rational numbers that tends to x. We may suppose that x_n lies between two fixed

rational numbers u and v for all values of n so that a^{x_n} lies between a^u and a^v and is therefore bounded, say $a^{x_n} < k$ for every value of n.

The sequence (a^{x_n}) tends to a limit, α say; because

$$|a^{x_{n+p}} - a^{x_n}| = |a^{x_n}(a^{x_{n+p}-x_n} - 1)| < k\,|a^{x_{n+p}-x_n} - 1|$$

and $|a^{x_{n+p}-x_n} - 1| \to 0$ if $n \to \infty$ since the sequence (x_n) is convergent and therefore $x_{n+p} - x_n \to 0$ if $n \to \infty$.

Again if (y_n) is any other sequence of rational numbers that tends to x the sequence (a^{y_n}) tends, by the above proof, to a limit, β say, but $\beta = \alpha$, as may be seen thus. The integer m may be chosen so large that x_n and y_n differ from x and from each other by less than ε if $n > m$; further,

$$|a^{y_n} - a^{x_n}| = |a^{x_n}(a^{y_n - x_n} - 1)| < k\,|a^{y_n - x_n} - 1|$$

and $|a^{y_n-x_n} - 1| \to 0$ when $n \to \infty$ because $y_n - x_n \to 0$ when $n \to \infty$. Hence $\beta = \alpha$.

The limit α is therefore the same whatever be the sequence (a^{x_n}) when (x_n) is a sequence of rational numbers that tends to x. α is defined to be the value of the function a^x for all real values of x, and it will be noticed that a^x cannot be negative.

(3) If $a > 1$ and c any rational number in the upper class of the section that defines x, when x is irrational, then $a^x < a^c$. Hence $a^x >$ or $< a^y$ according as $x >$ or $< y$, while if $a < 1$, $a^x >$ or $< a^y$ according as $x <$ or $> y$. These conclusions follow readily from § 11 and the definition of inequality (§ 5).

Again, it is easy to prove that the index laws

$$a^x \times a^y = a^{x+y}, \quad a^{-x} = 1/a^x, \ (a^x)^y = a^{xy}$$

are valid when x and y are any real numbers. Thus, for example, let x and y be defined by the sequences (x_n) and (y_n) of rational numbers; then, by the fundamental theorems of limits,

$$\mathop{\mathcal{L}}_{n\to\infty}(a^{x_n} \times a^{y_n}) = \mathop{\mathcal{L}}_{n\to\infty}(a^{x_n+y_n}), \text{ that is, } a^x \times a^y = a^{x+y}$$

since all the limits exist.

(4) a^x is a continuous function of x.

Suppose $a > 1$. It is always possible to choose *rational* numbers y and z such that, whether x is rational or irrational,

$$y < x < z \quad \text{and} \quad z - y = 1/n.$$

Now, let x' be any *real* number such that $y < x' < z$; then, by (3), $a^y < a^{x'} < a^z$. When y and z both tend to x so does x'; but in this case, by (2), both a^y and a^z tend to a^x and therefore $a^{x'}$ also tends to a^x.

A similar proof holds if $a < 1$. Hence, whether x' tends to x through rational or irrational values of x, $a^{x'}$ tends to a^x and therefore a^x is a continuous function of x.

(5) It follows now that, if $a > 1$, the function a^x is a monotonic, strictly increasing, continuous function of x; $a^x \to +\infty$ when $x \to +\infty$ and $a^x \to 0$ when $x \to -\infty$. It may further be noted that $a^x \to 1$ when $x \to 0$.

If $a < 1$ let $a = 1/b$ where $b > 1$ so that $a^x = 1/b^x$ and it follows that a^x is a monotonic, strictly decreasing function of x.

The student should note specially the following inequality, which is required in § 24, and is generally useful.

If $b > a > 1$ then $b^x > a^x$ when $x > 0$, but $b^x < a^x$ when $x < 0$, where x is a real number.

23. Logarithms. In this article a theorem will be assumed which is more conveniently discussed in the next chapter (§ 32), namely: If $f(x)$ is a monotonic, strictly increasing, continuous function of x for a given range of x the equation $f(x) = y$ defines x as a monotonic, strictly increasing, continuous function of y for the corresponding range of y. The two functions are said to be "inverse" to each other (*E.T.* p. 18).

In the preceding article it has been proved that a^x is, if $a > 1$, a monotonic, strictly increasing, continuous function of x that increases from 0 to ∞ as x increases from $-\infty$ to ∞. The inverse function is called a *logarithm*, and if $y > 0$ and $a^x = y$ the definition is that "x is the logarithm of y to the base a," so that, in the usual notation,

$$x = \log_a y \quad \text{if} \quad a^x = y, \quad \text{where } a > 1 \text{ and } y > 0.$$

It is obvious that $\log_a y = 0$ if $y = 1$ and that $\log_a y$ is positive or negative according as y is greater than or less than unity. (In practice the *base* a is usually assumed to be greater than unity.)

There is no need to discuss the well-known rules of operation with logarithms, but one property may be stated explicitly, namely, that if (y_n) is any sequence of real numbers that tends

to y the sequence $(\log y_n)$ tends to $\log y$. This property follows from the fact that $\log y$ is a continuous function of y (§ 21).

24. The Base e. Theorems in Limits. It has been proved in § 11, (10) and (10a) that, if n is a positive rational number, $(1+1/n)^n$ increases as n increases but that $(1+1/n)^n$ is less than 4 for every value of n. Therefore (§ 15, Th. I) $(1+1/n)^n$ tends to a limit, usually denoted by e, when $n \to \infty$.

Next, let z be any real number greater than unity and let the positive integer n be chosen so that $n \leqq z < n+1$; therefore when either of the numbers z and n tends to infinity so does the other.

Now
$$1+\frac{1}{n+1} < 1+\frac{1}{z} \leqq 1+\frac{1}{n},$$
and therefore
$$\left(1+\frac{1}{n+1}\right)^n < \left(1+\frac{1}{z}\right)^z < \left(1+\frac{1}{n}\right)^{n+1}.$$
But
$$\left(1+\frac{1}{n+1}\right)^n = \left(1+\frac{1}{n+1}\right)^{n+1} \Big/ \left(1+\frac{1}{n+1}\right),$$
$$\left(1+\frac{1}{n}\right)^{n+1} = \left(1+\frac{1}{n}\right)^n \left(1+\frac{1}{n}\right),$$
and therefore both $\left(1+\frac{1}{n+1}\right)^n$ and $\left(1+\frac{1}{n}\right)^{n+1}$ tend to e when $n \to \infty$, since $\left(1+\frac{1}{n}\right)^n$ does so for all rational values of n. Hence $1+\frac{1}{z}\Big)^z$ tends to e when z tends to infinity through real (positive) values.

Next let $z = -y-1$ where $y > 0$; then $z \to -\infty$ when $y \to +\infty$.

But
$$1+\frac{1}{z} = \left(\frac{y}{y+1}\right)^{-y-1} = \left(1+\frac{1}{y}\right)^y \Big(1+\frac{1}{y}$$
and therefore $\left(1+\frac{1}{z}\right)^z$ tends to e when $z \to -\infty$.

The above result may be stated as a theorem.

THEOREM I. *The expression* $\left(1+\frac{1}{z}\right)^z$ *tends to a definite limit, denoted by e, when z tends through real values either to* $+\infty$ *or to* $-\infty$; *or, (if* $1/z$ *is substituted for z) the expression* $(1+z)^{1/z}$ *tends to the limit e when z tends to zero through real values which may be either positive or negative.*

This number e, as the student knows, is taken as the base of the logarithm in all theoretical investigations, and the function e^x is "the exponential function of x." If a is positive the function a^x, which is of course an exponential function of x, may be expressed, in what is considered the standard form, as $e^{x \log a}$. In accordance with the usual practice, the symbol $\log a$ is here taken to mean $\log_e a$.

THEOREM II. $\displaystyle \mathop{\mathcal{L}}_{h\to 0} \frac{a^h-1}{h} = \log a,\ h>0 \text{ or } h<0.$

Suppose $a>1$. Then $a^h = 1+k$; k is positive or negative according as h is positive or negative and $k\to 0$ when $h\to 0$. Now $h\log a = \log(1+k)$, and therefore

$$\frac{a^h-1}{h} = \frac{k\log a}{\log(1+k)} = \frac{\log a}{\log[(1+k)^{1/k}]}.$$

But (Th. I) $(1+k)^{1/k}\to e$ and therefore, since $\log x$ is a continuous function of x, $\log[(1+k)^{1/k}]\to \log e$ when $k\to 0$. The theorem is therefore proved since $\log e = 1$.

If $0<a<1$, let $b=1/a$; then $b>1$ so that

$$\mathop{\mathcal{L}}_{h\to 0} \frac{a^h-1}{h} = -\mathop{\mathcal{L}}_{h\to 0} \frac{1}{b^h}\frac{b^h-1}{h} = -\log b = \log a.$$

Cor. Let $h=1/n$ where n is a positive integer; then $n(\sqrt[n]{a}-1)$ tends to $\log a$ when n tends to infinity.

THEOREM III. $\displaystyle \mathop{\mathcal{L}}_{h\to 0} \frac{\log(x+h)-\log x}{h} = \frac{1}{x}.$

Since x must be positive we suppose $x \geqq c > 0$. Now

$$\frac{\log(x+h)-\log x}{h} = \frac{1}{h}\log\left(1+\frac{h}{x}\right) = \frac{1}{x}\log\left[\left(1+\frac{h}{x}\right)^{\frac{x}{h}}\right].$$

Let $h/x = k$; then $k\to 0$ when $h\to 0$ since $x\geqq c>0$. As before $\log[(1+k)^{1/k}]$ tends to $\log e$, that is, to unity when $k\to 0$, so that the theorem is established.

Theorems II and III show at once that the derivatives of e^x, a^x and $\log x$ are e^x, $a^x \log a$ and $1/x$ respectively.

Derivative of x^n when n is irrational. Since x^n, when $x>0$, may be expressed as $e^{n\log x}$ we find

$$\frac{d\,.\,x^n}{dx} = \frac{d\,.\,e^{n\log x}}{dx} = e^{n\log x}\,.\,\frac{d(n\log x)}{dx} = x^n\,.\,\frac{n}{x} = nx^{n-1},$$

so that the usual rule for the derivative of x^n applies whether n is rational or irrational. However, when n is irrational, x must be positive.

Notation. The exponential function e^x is frequently expressed by the notation $\exp(x)$; this notation is specially useful when the index is a somewhat lengthy expression such as $(ax^2+bx+c)/(a'x^2+b'x+c')$ or $n\log(1+x/n)$.

25. Limits and Inequalities. The examples now to be given establish some theorems in limits of logarithmic and exponential functions and also some inequalities that are frequently required.

Ex. 1. If the product nk tends to x when $|n|$ tends to infinity, so does the product $n\log(1+k)$.

The proof depends on the theorem which has been applied so often in § 24 that $\log[(1+k)^{1/k}]$ tends to $\log e$, that is, to unity when k tends in any way to zero.

If nk tends to x whether n tends to $+\infty$ or to $-\infty$ then k tends to zero whether x is or is not zero. Now

$$n\log(1+k)=nk\log[(1+k)^{1/k}]$$

and therefore

$$\mathop{\mathcal{L}}_{|n|\to\infty} n\log(1+k)=\mathop{\mathcal{L}}_{|n|\to\infty}(nk)\,.\mathop{\mathcal{L}}_{k\to 0}\log[(1+k)^{1/k}]=x.$$

Ex. 2. If the product nk tends to x when $|n|$ tends to infinity, then $(1+k)^n$ tends to e^x.

$$(1+k)^n=e^{x_n} \text{ where } x_n=n\log(1+k).$$

By Ex. 1, $x_n\to x$ and therefore, by § 22, (4), $e^{x_n}\to e^x$.

Two particular cases of this theorem are important.

(i) Let $k=x/n$; then $(1+x/n)^n$ tends to e^x when n tends either to $+\infty$ or to $-\infty$.

(ii) Let $nx=m$; then, if x is not zero, $|m|$ tends to infinity when $|n|$ does so. Therefore

$$\left(1+\frac{1}{n}\right)^{nx}=\left(1+\frac{x}{m}\right)^m \text{ so that } \left(1+\frac{1}{n}\right)^{nx}\to e^x.$$

If $x=0$, $(1+1/n)^{nx}=1$, and therefore the limit is 1 and $1=e^0$.

In these and similar examples it has to be remembered that n and k must be such that $(1+k)$ is *positive.*

Ex. 3. Prove the following inequalities:

(i) $e^x>1+x$ if $x>0$; (ii) $e^{-x}>1-x$ if $0<x<1$;

(iii) $x-\frac{1}{2}x^2<\log(1+x)<x$ if $0<x<1$;

(iv) $x < -\log(1-x) < x + \frac{1}{2}\frac{x^2}{1-x}$ if $0 < x < 1$;

(v) $2x < \log\frac{1+x}{1-x} < 2x + \frac{2}{3}\frac{x^3}{1-x^2}$ if $0 < x < 1$.

For the extreme values 0 and 1 of x it is best to test the value of the function. In all cases the inequalities become equalities for $x=0$; for $x=1$ the inequalities (ii) and (iii) persist but the logarithms in (iv) and (v) tend to infinity when $x \to 1$.

These results are proved very simply by the method used in the *Elementary Treatise*, p. 132, examples 24, 33, 34. To illustrate the method take the inequalities (v). Let

$$f(x) = \log\frac{1+x}{1-x} - 2x;\quad \varphi(x) = \log\frac{1+x}{1-x} - \left(2x + \frac{2}{3}\frac{x^3}{1-x^2}\right).$$

Here $$f'(x) = \frac{2x^2}{1-x^2},\quad \varphi'(x) = -\frac{4}{3}\frac{x^4}{(1-x^2)^2},$$

and therefore $f'(x)$ is positive and $\varphi'(x)$ negative if $0 < x < 1$. But $f(x)$ and $\varphi(x)$ are both zero when $x=0$ so that the increasing function $f(x)$ is positive and the decreasing function $\varphi(x)$ is negative for the range $0 < x < 1$. The inequalities are therefore proved.

Ex. 4. Prove that $$e < \left(1+\frac{1}{n}\right)^{n+\frac{1}{2}} < e \times e^{\frac{1}{12n(n+1)}}$$

where n is any positive number.

The inequalities (v) of Ex. 3 may be expressed in the form

$$1 < \frac{1}{2x}\log\frac{1+x}{1-x} < 1 + \frac{1}{3}\frac{x^2}{1-x^2},\quad 0 < x < 1.$$

Now let $x = 1/(2n+1)$ where $n > 0$; then

$$\frac{1}{2x}\log\frac{1+x}{1-x} = \log\left[\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}}\right],\quad \frac{1}{3}\frac{x^2}{1-x^2} = \frac{1}{12n(n+1)},$$

so that $$1 < \log\left[\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}}\right] < 1 + \frac{1}{12n(n+1)},$$

and, therefore, passing from logarithms to numbers, we find the inequalities stated in the example.

Ex. 5. If $\varphi(n) = (\cos\theta_n)^{f(n)}$ and if $\theta_n \to 0$ and $f(n) \to \infty$ when $n \to \infty$, discuss the question of a limit for $\varphi(n)$ when $n \to \infty$.

Let $u_n = f(n)\log(\cos\theta_n)$; then $\varphi(n) = e^{u_n}$ and the problem reduces to that of finding the limit of u_n. This problem is, however, indeterminate until a relation between θ_n and $f(n)$ is given.

Now by Ex. 3, (iv) the limit of $[\log(1-x)]/x$ when x tends to zero is -1. But $\log(\cos\theta_n) = \frac{1}{2}\log(1-\sin^2\theta_n)$ and

$$u_n = \tfrac{1}{2}f(n)\log(1-\sin^2\theta_n) = \tfrac{1}{2}f(n)\,.\,\theta_n^2\left[\frac{\log(1-\sin^2\theta_n)}{\sin^2\theta_n}\right]\cdot\left(\frac{\sin\theta_n}{\theta_n}\right)^2,$$

so that the limit of u_n is the same as the limit of $-\frac{1}{2}f(n)\theta_n^2$ since the

factors $\sin\theta_n/\theta_n$ and $\log(1-\sin^2\theta_n)/\sin^2\theta_n$ tend to 1 and -1 respectively. Hence

$$\text{if } f(n)\,\theta_n^2 \to 0,\ u_n \to 0,\ \varphi(n) \to 1;$$
$$\text{if } f(n)\,\theta_n^2 \to a^2,\ u_n \to -\tfrac{1}{2}a^2,\ \varphi(n) \to e^{-\frac{1}{2}a^2};$$
$$\text{if } f(n)\,\theta_n^2 \to +\infty,\ u_n \to -\infty,\ \varphi(n) \to 0.$$

If $\varphi(n) = (\sin\theta_n/\theta_n)^{f(n)}$ let $\sin\theta_n/\theta_n = 1 - \dfrac{\theta_n^2 + \alpha}{6}$, where $\alpha \to 0$ when $\theta_n \to 0$ (that is, use the first two terms of the series for $\sin\theta_n$) and then proceed as before. *Or*, note that (*E.T.* p. 77) $\sin\theta_n/\theta_n$ lies between 1 and $\cos\theta_n$ when $|\theta_n| < \dfrac{\pi}{2}$.

Ex. 6. If $m > 0$, $(\log x)/x^m \to 0$ when $x \to \infty$.

Let $x = e^y$ and apply the method of Examples 8, 9, p. 99 of the *Elementary Treatise.*

26. Extension of Range of Definition. The method by which the range of definition of the function a^x has been extended from rational to real values of x is of a general character, and may often be applied. The principle of the method may be stated in the following way.

Suppose that a function $f(x)$ has the two properties:

(1) $f(x)$ is a bounded, single-valued function of x, defined for all *rational* values of x in the range $a \leqq x \leqq b$;

(2) if ξ is any rational value of x in the range, $f(x)$ tends to $f(\xi)$ when x tends through rational values to ξ.

Now *every* point ξ in the interval (a, b) is a limiting point of the set of rational points in the interval and by Cauchy's Test (§ 21) $f(x)$ tends to a limit, l_ξ say, when x tends through rational values to ξ whether ξ is rational or irrational. When ξ is rational l_ξ is the value $f(\xi)$ of the function; now, if ξ is irrational, let l_ξ be defined to be *the value* of $f(x)$ when $x = \xi$ and $f(x)$ will be defined for *all* values of x in (a, b), and further, $f(x)$ will be *continuous* for the range $a \leqq x \leqq b$.

To see that $f(x)$ is continuous, let ξ and c be two values of x in the range, ξ being rational or irrational and c rational. When x tends to ξ through *rational* values, $f(x)$, by the extended definition, tends to $f(\xi)$; it has to be proved that $f(x)$ tends to $f(\xi)$ whether x tends to ξ through rational or through irrational values.

In the first place, η may be chosen so that

$$|f(c) - f(\xi)| < \tfrac{1}{2}\varepsilon \quad \text{if} \quad |c - \xi| < \eta \quad \text{..............(i)}$$

Next, let ξ' be any number, rational or irrational, *between* ξ and c; there is in the interval (ξ', c) a *rational* number c' such that $|f(c')-f(\xi')|<\frac{1}{2}\varepsilon$. It is possible that c' may be taken to be c; if not, substitute c' for c in the inequality (i) (this substitution is manifestly admissible) and let the one symbol c' be used to cover both cases. We now have

$$|c'-\xi'|<|c'-\xi|<\eta, \quad |\xi'-\xi|<\eta$$

and $$|f(c')-f(\xi)|<\tfrac{1}{2}\varepsilon, \quad |f(c')-f(\xi')|<\tfrac{1}{2}\varepsilon,$$

so that, if $|\xi'-\xi|<\eta$, we find

$$|f(\xi')-f(\xi)|=|\{f(c')-f(\xi)\}-\{f(c')-f(\xi')\}|$$
$$<\tfrac{1}{2}\varepsilon+\tfrac{1}{2}\varepsilon, \text{ or, } \varepsilon.$$

Thus $f(x)$ tends to $f(\xi)$ when x tends to ξ through irrational as well as through rational values, and therefore $f(x)$ is continuous at ξ.

EXERCISES II.

1. If $$a_n=\frac{1}{\sqrt{(n^2+1)}}+\frac{1}{\sqrt{(n^2+2)}}+\frac{1}{\sqrt{(n^2+3)}}+\dots+\frac{1}{\sqrt{(n^2+n)}}$$
$a_n \to 1$ when $n \to \infty$.

2. If $$a_n=\frac{n}{n^2+1^2}+\frac{n}{n^2+2^2}+\frac{n}{n^2+3^2}+\dots+\frac{n}{n^2+n^2}$$
$a_n \to \pi/4$ when $n \to \infty$.

3. If $a_{n+1}=\frac{1}{2}(a_n+b_n)$ and $b_{n+1}=\sqrt{(a_{n+1}b_n)}$, $a_n>0$, $b_n>0$, the sequences (a_n) and (b_n) are monotonic and converge to the same limit, α say.

If $a_1=\cos\theta$, $b_1=1$, then $\alpha=\sin\theta/\theta$; if $a_1=\cosh u$, $b_1=1$ then $\alpha=\sinh u/u$. (Borchardt.)

4. If $a_{n+2}=\sqrt{(a_{n+1}a_n)}$ and $a_n>0$, the sequences (a_{2n-1}) and (a_{2n}) are both monotonic, one increasing and the other decreasing; the sequence (a_n) tends to $(a_1a_2^2)^{\frac{1}{3}}$.

5. If $a_{n+2}=\frac{1}{2}(a_{n+1}+a_n)$ and $a_n>0$, the sequences (a_{2n-1}) and (a_{2n}) behave as in Ex. 4, and the sequence (a_n) tends to $\frac{1}{3}(a_1+2a_2)$.

6. If $a_{n+1}=\sqrt{(a+a_n)}$, where each number is positive, the sequence (a_n) tends to ξ, where ξ is the positive root of the equation $x^2=x+a$.

[Here $a_{n+1}^2-a_{n+2}^2=a_n-a_{n+1}$ so that (each number being positive) $a_{n+1}>$ or $<a_{n+2}$ according as $a_n>$ or $<a_{n+1}$, that is, (a_n) is monotonic. Again, since $\xi^2=\xi+a$ we have $a_2^2-\xi^2=a_1-\xi$ so that $a_2>\xi$ if $a_1>\xi$; also if $a_1>\xi'$ (the positive root of $x^2=x+a$) then $a_1^2>a+a_1=a_2^2$ or $a_1>a_2$. Thus when (a_n) is a decreasing sequence each a_n is greater than ξ, with a similar conclusion if $a_1<\xi$. Hence, since a_n and a_{n+1}

tend to the same (positive) limit, η say, we have $\eta^2 = a + \eta$ so that $\eta = \xi$.]

7. If $a_n = a/(1 + a_{n-1})$ (each number positive) the sequence (a_n) tends to a limit ξ, the positive root of the equation $x^2 + x = a$.

8. If $u_n = H_n - \log n$ and $v_n = H_n - \log(n+1)$ where

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n},$$

show that (u_n) and (v_n) tend to the same limit, γ say; γ is a common (though not universal) notation for the limit which is always known as **Euler's Constant.** $(\gamma = 0{\cdot}577\ 215\ 664\ \ldots)$.

[By the inequalities, § 25, Ex. 3, we have

$$u_n - u_{n+1} = \frac{-1}{n+1} - \log\left(1 - \frac{1}{n+1}\right) > 0;$$

$$v_n - v_{n+1} = \frac{-1}{n+1} + \log\left(1 + \frac{1}{n+1}\right) < 0;$$

$$u_n - v_n = \log\left(1 + \frac{1}{n}\right) > 0,\ u_n > v_n.$$

Thus (u_n) is decreasing, (v_n) increasing, but $u_n > v_1$ and $v_n < u_1$. Hence $(u_n) \to$ a limit α and $v_n \to$ a limit β while $(u_n - v_n) \to 0$, so that $\alpha = \beta = \gamma$.

Or, we may proceed as follows:

$$n + 1 = \frac{2}{1} \cdot \frac{3}{2} \cdot \ldots \frac{n+1}{n} = \left(1 + \frac{1}{1}\right)\left(1 + \frac{1}{2}\right) \ldots \left(1 + \frac{1}{n}\right),$$

so that
$$v_n = \sum_{r=1}^{n} \left\{\frac{1}{r} - \log\left(1 + \frac{1}{r}\right)\right\};\ 0 < v_n < \tfrac{1}{2} \sum_{r=1}^{n} \frac{1}{r^2}.$$

The series $\sum 1/r^2$ converges and therefore $v_n \to$ a limit; $u_n - v_n \to 0$ and the sequence (u_n) tends to the same limit.]

Cor. $H_n = \gamma + \log n + \theta_n$ where $\theta_n \to 0$ when $n \to \infty$. This expression for the sum of the first n terms of the harmonic series is often useful.

9. If $a_n = \dfrac{1}{n+1} + \dfrac{1}{n+2} + \ldots + \dfrac{1}{n+n}$,

prove that $a_n \to \log 2$ when $n \to \infty$ and deduce a series for $\log 2$.

$$\Big[a_n = H_{2n} - H_n = (\gamma + \log 2n + \theta_n') - (\gamma + \log n + \theta_n''),$$

so that
$$a_n = \log 2 + (\theta_n' - \theta_n'') \to \log 2 \text{ since } \theta_n' \to 0,\ \theta_n'' \to 0.$$

Again
$$H_{2n} - H_n = \sum_{r=1}^{2n} \frac{1}{r} - 2\sum_{r=1}^{n} \frac{1}{2r} = \sum_{r=1}^{2n} \frac{(-1)^{r-1}}{r},$$

so that
$$\log 2 = \sum_{r=1}^{\infty} \frac{(-1)^{r-1}}{r} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots.\Big]$$

10. If
$$a_n = \sum_{r=1}^{2n} \frac{1}{2r-1} \text{ and } b_n = \sum_{r=1}^{n} \frac{1}{2r},$$

prove that $$\mathop{\mathcal{L}}(a_n - b_n) = \tfrac{3}{2}\log 2,$$

and deduce that

$$\mathop{\mathcal{L}}\left(1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \dots + \frac{1}{4n-3} + \frac{1}{4n-1} - \frac{1}{2n}\right) = \frac{3}{2}\log 2.$$

[Here $a_n - b_n = H_{4n} - \frac{1}{2}H_{2n} - \frac{1}{2}H_n$. The infinite series

$$1 + \tfrac{1}{3} - \tfrac{1}{2} + \tfrac{1}{5} + \tfrac{1}{7} - \tfrac{1}{4} + \dots$$

is a derangement* of the series in Ex. 9 for log 2, so that this derangement alters the sum of the series.]

11. If p and q are fixed positive integers and if

$$c_r = \sum_{s=1}^{p} \frac{1}{2(r-1)p + 2s - 1}, \quad d_r = \sum_{s=1}^{q} \frac{1}{2(r-1)q + 2s},$$

$$a_n = \sum_{r=1}^{n} c_r, \quad b_n = \sum_{r=1}^{n} d_r,$$

show that

$$\text{(i)} \quad a_n - b_n = H_{2pn} - \tfrac{1}{2}H_{pn} - \tfrac{1}{2}H_{qn}.$$

$$\text{(ii)} \quad \mathop{\mathcal{L}}(a_n - b_n) = \log 2 + \frac{1}{2}\log\frac{p}{q}.$$

State the result as a theorem on the change produced in the value of the series for log 2 (Ex. 9) by a derangement* of its terms.

12. If p is a fixed positive integer and if

$$a_n = \frac{1}{n+1} + \frac{1}{n+2} + \frac{1}{n+3} + \dots + \frac{1}{n+pn},$$

and

$$b_n = \frac{2}{2n+1} + \frac{2}{2n+3} + \frac{2}{2n+5} + \dots + \frac{2}{2n+2pn-1},$$

show that both a_n and b_n tend to $\log(p+1)$ when $n \to \infty$.

13. If $a_n = \{\frac{1}{2}(a^{1/n} + b^{1/n})\}^n$, show that a_n tends to $\sqrt{(ab)}$ when $n \to \infty$.
[Here $a_n = (1+k)^n$ where $nk = \frac{1}{2}\{n(a^{1/n} - 1) + n(b^{1/n} - 1)\}$ and (§ 24, Th. II, Cor.) $nk \to \frac{1}{2}\log(ab)$ when $n \to \infty$.]

14. $$\mathop{\mathcal{L}}_{n\to\infty} \frac{1^p + 2^p + 3^p + \dots + n^p}{n^{p+1}} = \frac{1}{p+1}, \quad p+1 > 0.$$

15. $$\mathop{\mathcal{L}}_{n\to\infty} \left(\frac{1^p + 2^p + 3^p + \dots + n^p}{n^p} - \frac{n}{p+1}\right) = \frac{1}{2}, \quad p > 0.$$

16. $$\mathop{\mathcal{L}}_{n\to\infty} \left\{\frac{(a+c)^p + (a+2c)^p + \dots + (a+nc)^p}{n^p} - \frac{nc^p}{p+1}\right\}$$
$$= (a + \tfrac{1}{2}c)c^{p-1}, \quad p > 0, \ a \geqq 0, \quad c > 0.$$

* On derangement of series, see § 59.

17. If $f(n)=\sum_{s=1}^{\infty}\left(\frac{n}{n+s}\right)^{1+\rho}$, $\rho>0$, show that $n^{-1-\rho}f(n)\to 0$ when $n\to\infty$ and then prove

$$\text{(i)}\ \lim_{n\to\infty} n^{-1}f(n)=\frac{1}{\rho};\quad \text{(ii)}\ \lim_{n\to\infty}\left\{\frac{n}{\rho}-f(n)\right\}=\frac{1}{2}.$$

18. $\lim \sqrt[n]{n}=\lim \frac{n}{n-1}=1.$ (Cesàro).

19. $\lim_{n\to\infty}\frac{1+\sqrt{2}+\sqrt[3]{3}+\sqrt[4]{4}+\ldots+\sqrt[n]{n}}{n}=1.$

20. $\lim_{n\to\infty}\frac{1}{n}\sqrt[n]{(n!)}=\lim_{n\to\infty}\sqrt[n]{\left(\frac{n!}{n^n}\right)}=\frac{1}{e}.$

21. $\lim_{n\to\infty}\frac{1}{n}\left\{(n+1)(n+2)\ldots(n+n)\right\}^{\frac{1}{n}}=\frac{4}{e}.$

22. $\lim \frac{1}{n}\left\{(k+1)(k+2)\ldots(k+n)\right\}^{\frac{1}{n}}=\frac{1}{e}$, ($k$ fixed).

23. If $a_n=c_0x_0+c_1x_1+c_2x_2+\ldots+c_nx_n$, $b_n=c_0+c_1+c_2+\ldots+c_n$, where $c_r(r=0, 1, 2, \ldots, n)$ is positive (at least for all values of r greater than a fixed integer) and $b_n\to\infty$ when $n\to\infty$, prove that if x_n tends to a limit l

$$\lim_{n\to\infty}\frac{a_n}{b_n}=\lim_{n\to\infty}\frac{c_0x_0+c_1x_1+\ldots+c_nx_n}{c_0+c_1+\ldots+c_n}=l.$$

l is not necessarily finite (see § 20, Ex. 2).

24. If the symbol $\binom{n}{r}$ denotes the binomial coefficient

$$n(n-1)(n-2)\ldots(n-r+1)/r!$$

show that, when r is a fixed integer, $\frac{1}{2^n}\binom{n}{r}\to 0$ when $n\to\infty$.

[Here $\qquad \frac{1}{2^n}\binom{n}{r}<\frac{1}{r!}\cdot\frac{n^r}{2^n}=\frac{1}{r!}\frac{n^r}{e^{n\log 2}}\to 0$ if $n\to\infty$.]

25. If $a_n=x_0+\binom{n}{1}x_1+\binom{n}{2}x_2+\ldots+\binom{n}{r}x_r+\ldots+\binom{n}{n}x_n$, show that $a_n/2^n\to 0$ if $(x_n)\to 0$, and that $a_n/2^n\to x$ if $(x_n)\to x$.

$$\left[1+\binom{n}{1}+\binom{n}{2}+\ldots+\binom{n}{r}+\ldots+\binom{n}{n}=(1+1)^n=2^n,\right.$$

so that if $r>1$, $\frac{1}{2^n}\left[\binom{n}{r+1}+\binom{n}{r+2}+\ldots+\binom{n}{n}\right]<1.$

Now if $(x_n)\to 0$ we can choose r so that $|x_p|<\frac{1}{2}\epsilon$ if $p>r$ and then, if $n>r$,

$$\frac{1}{2^n}\left[\binom{n}{r+1}x_{r+1}+\binom{n}{r+2}x_{r+2}+\ldots+\binom{n}{n}x_n\right]<\frac{1}{2}\epsilon.$$

Choose r and keep it fixed; the other part of $a_n/2^n$ contains a *finite* number of terms, namely $(r+1)$, each of which tends to zero when $n \to \infty$, and therefore their sum will be less than $\frac{1}{2}\varepsilon$ when $n > N$. Thus $a_n/2^n$ is less than ε if $n > N$ so that $a_n/2^n \to 0$ if $x_n \to 0$.

If (x_n) tend to x we may write

$$x = \frac{1}{2^n}\sum_{r=0}^{n}\binom{n}{r}x, \quad b_n = \sum_{r=0}^{n}\binom{n}{r}(x_r - x),$$

and $b_n/2^n \to 0$ since $(x_n - x) \to 0$ so that $a_n/2^n \to x.$]

26. $\mathop{\mathcal{L}}\limits_{x\to 0}(\cos x)^{\frac{1}{x^2}} = e^{-\frac{1}{2}}$ and $\mathop{\mathcal{L}}\limits_{x\to 0}(\sin x)^{\sin x} = 1.$

27. If $\tan \alpha_k = \lambda \tan (k\pi/n)$, where $0 < \lambda \leqq 1$ and k is a fixed positive integer, prove that

$$\text{(i)} \mathop{\mathcal{L}}\limits_{n\to\infty} \cos (n\alpha_k) = \cos \lambda k\pi; \quad \text{(ii)} \mathop{\mathcal{L}}\limits_{n\to\infty} (\cos \alpha_k)^n = 1.$$

The following Examples 28-30 lead to Stirling's approximation for $n!$

28. If $\varphi(n) = n!\,e^n/n^{n+\frac{1}{2}}$ apply Ex. 4 of § 25 to prove that $\varphi(n)$ is a monotonic, decreasing, positive function of n and therefore tends to a limit k when $n \to \infty$. Show that $k > 0$.

$$\left[\frac{\varphi(n+1)}{\varphi(n)} = \frac{e}{(1+1/n)^{n+\frac{1}{2}}} < 1; \quad \varphi(n+1) < \varphi(n).\right.$$

To prove $k > 0$, let $\psi(n) = e^{-\frac{1}{12n}}\varphi(n)$; then $\psi(n+1) > \psi(n)$. Since $\psi(n)$ has the same limit as $\varphi(n)$ and increases to its limit k, it follows that $k > 0.$]

29. From Wallis's expression for $\pi/2$ (*E.T.* p. 307) prove that $\sqrt{(\pi/2)}$ lies between $P_n/\sqrt{(2n)}$ and $P_n/\sqrt{(2n+1)}$ where

$$P_n = (2^n \,.\, n!)^2/(2n)!$$

and therefore

$$(2^n \,.\, n!)^2/(2n)! = \sqrt{\left(\frac{\pi}{2}\right)}\sqrt{(2n+\theta_n)}, \quad 0 < \theta_n < 1.$$

Deduce that

$$\frac{1\,.\,3\,.\,5 \ldots (2n-1)}{2\,.\,4\,.\,6 \ldots (2n)} = \frac{a_n}{\sqrt{n}}, \text{ where } a_n \to \frac{1}{\sqrt{\pi}} \text{ when } n \to \infty.$$

30. Show that $\dfrac{[\varphi(n)]^2}{\varphi(2n)} = \dfrac{\sqrt{(2n+\theta_n)}}{\sqrt{n}}\,.\,\sqrt{\pi},\ 0 < \theta_n < 1,$

and therefore that $k = \sqrt{(2\pi)}$, where k is defined in Ex. 28, so that $\varphi(n) \to \sqrt{(2\pi)}$ when $n \to \infty$.

[By Ex. 28, $\varphi(n)$ is greater than k and $\psi(n)$ or $e^{-\frac{1}{12n}}\varphi(n)$ is less than k or $\sqrt{(2\pi)}$; that is

$$\varphi(n) > \sqrt{(2\pi)} \quad \text{but } \varphi(n) < \sqrt{(2\pi)}e^{\frac{1}{12n}}$$

and therefore $$\varphi(n) = \sqrt{(2\pi)}e^{\theta/12n},\ 0<\theta<1.$$

Thus finally $$n! = \left(\frac{n}{e}\right)^n \sqrt{(2\pi n)}e^{\theta/12n}, \quad 0<\theta<1.$$

The factor $e^{\theta/12n}$ is less than $1+\frac{1}{11n}(n>1)$ and tends to unity when $n\to\infty$. The value

$$(n/e)^n\sqrt{(2\pi n)}$$

is known as **Stirling's Approximation** to $n!$ when n is large.

For the form in which Stirling states his theorem see Tweedie's *James Stirling: A Sketch of his Life and Works*, pp. 43-44.]

function $f(x)$ is not bounded above. Similarly, if $f(x)$ is not bounded below in (a, b), there is a point ξ' in (a, b) such that in the interval $(\xi' - \varepsilon, \xi' + \varepsilon)$ the function $f(x)$ is not bounded below.

Note. A function $f(x)$ may be finite for every given value of x in the closed range (a, b) and yet not bounded in (a, b). For example, let $f(x)$ be defined as the limit when $n \to \infty$ of $nx/(1+nx^2)$; then $f(x)=0$ if $x=0$, but $f(x)=1/x$ if x is not zero. Thus $f(x)$ is finite for every given value of x but, in any interval which contains the value 0 of x, the function $f(x)$ is not bounded since, if K is any arbitrarily large positive number, $|f(x)| > K$ when $0 < |x| < 1/K$.

28. Theorems on Continuous Functions. Throughout this article the function $f(x)$ is supposed to be single-valued and continuous for a range $a \leqq x \leqq b$, or in the closed interval (a, b); in the interval x varies continuously—that is, x may take any value between a and b, including a and b.

The phrase "neighbourhood of ξ" will be used occasionally, and by a neighbourhood is meant the set of values of x in the interval $(\xi - \delta, \xi + \delta)$, excluding ξ, where δ is an arbitrarily small positive number. If $\xi = a$ the interval is $(a, a+\delta)$ and if $\xi = b$ the interval is $(b - \delta, b)$.

THEOREM I. *If $f(x)$ is continuous at c and if $f(c)$ is not zero, then $f(x)$ has the same sign as $f(c)$ for all values of x in the neighbourhood of c.*

By the definition of continuity $|f(x) - f(c)| < \varepsilon$ if $|x - c| < h$ so that $f(c) - \varepsilon < f(x) < f(c) + \varepsilon$ if $0 < |x - c| < h$. When $f(c)$ is not zero ε may be chosen so that both $f(c) - \varepsilon$ and $f(c) + \varepsilon$ have the same sign as $f(c)$, and therefore $f(x)$ has the same sign as $f(c)$ when $0 < |x - c| < h$.

THEOREM II. *If $f(x)$ is continuous for the range $a \leqq x \leqq b$ and if $f(a)$ and $f(b)$ have opposite signs, $f(x)$ will be zero for at least one value of x between a and b; further, if $f(a) = A$ and $f(b) = B$, $f(x)$ will take once at least every value between A and B when x varies continuously from a to b.*

The second part of the theorem is a simple corollary of the first part. For, if $A < C < B$ or $A > C > B$, let $\varphi(x) = f(x) - C$;

then $\varphi(x)$ is continuous for $a \leqq x \leqq b$, $\varphi(a) = A - C$, $\varphi(b) = B - C$ so that $\varphi(a)$ and $\varphi(b)$ have opposite signs. Therefore, by the first part, there is at least one value ξ, where $a < \xi < b$, such that $\varphi(\xi) = 0$ and therefore $f(\xi) = C$.

To prove the first part of the theorem suppose, for definiteness, that $f(a)$ is negative and $f(b)$ positive, and apply the method of the decreasing interval.

First let $c = \frac{1}{2}(a + b)$. If $f(c) = 0$ the theorem is proved, but if $f(c)$ is not zero, let $a = a_1$ and $c = b_1$ when $f(c)$ is *positive*, but let $c = a_1$ and $b = b_1$ when $f(c)$ is *negative*. Thus $f(a_1)$ is negative, $f(b_1)$ is positive and $f(x)$ is continuous for $a_1 \leqq x \leqq b_1$ where $b_1 - a_1$ is equal to $\frac{1}{2}(b - a)$.

Now repeat this process. If $c_1 = \frac{1}{2}(a_1 + b_1)$ either $f(c_1) = 0$, in which case the theorem is proved, or else $f(c_1)$ is not zero, and then we take $a_1 = a_2$, $c_1 = b_2$ when $f(c_1)$ is positive, but $c_1 = a_2$, $b_1 = b_2$ when $f(c_1)$ is negative. Hence $f(a_2)$ is negative, $f(b_2)$ positive and $f(x)$ is continuous for $a_2 \leqq x \leqq b_2$, while

$$b_2 - a_2 = \tfrac{1}{2}(b_1 - a_1) = \frac{1}{2^2}(b - a).$$

Proceeding in this way we find *either* a number, c_r say, for which $f(c_r) = 0$, in which case the theorem is proved, *or else* a sequence (a_n, b_n) of intervals which determines a number ξ, common to every interval, and $f(a_n)$ is negative, $f(b_n)$ positive for every value of n.

The continuity of $f(x)$ now comes into play. If $f(\xi)$ is not zero $f(x)$ has the same sign as $f(\xi)$ in the neighbourhood of ξ. But however small the positive number h may be, n may be chosen so that the interval (a_n, b_n) lies wholly within the interval $(\xi - h, \xi + h)$ and therefore, since $f(a_n)$ and $f(b_n)$ have opposite signs, $f(x)$ has not always the same sign as $f(\xi)$ when x lies in $(\xi - h, \xi + h)$. Hence $f(\xi)$ must be zero and the theorem is proved.

Ex. If n is a positive integer and k a positive (real) number the equation $x^n = k$ has one and only one positive (real) root.

Take b so that $b > k$ and also $b > 1$. Then $x^n = 0$ if $x = 0$ and $x^n = b^n > k$ if $x = b$. Therefore as x varies from 0 to b the continuous function x^n must, for at least one value of x, be equal to k. Further, if $x > 0$ and $y > 0$, x^n and y^n are unequal if x and y are unequal, so that there is only one positive value of x that makes $x^n = k$.

THEOREM III. *If $f(x)$ is continuous for the range $a \leqq x \leqq b$ and if ε is any given arbitrarily small positive number there is a positive number h such that $|f(x') - f(x'')| < \varepsilon$, where x' and x'' are any two values of x in the range such that $|x' - x''| < h$.*

Several proofs of this important theorem have been given; the following is by Peano.

First, choose $a_1 > a$ so that $|f(x) - f(a)| < \frac{1}{3}\varepsilon$ if $a \leqq x \leqq a_1$; this choice is possible because of the continuity of $f(x)$. Next choose $a_2 > a_1$ so that $|f(x) - f(a_1)| < \frac{1}{3}\varepsilon$ if $a_1 \leqq x \leqq a_2$, and let this process be continued. It has to be proved that a *finite* number of values, say $a_1, a_2, a_3, \ldots, a_n$, can be found such that in each of the $(n+1)$ intervals

$$(a, a_1), (a_1, a_2) \ldots, (a_{n-1}, a_n), (a_n, b) \ldots\ldots\ldots\ldots\ldots(1)$$

$$|f(x) - f(a_r)| < \tfrac{1}{3}\varepsilon \text{ if } a_r \leqq x \leqq a_{r+1}, \ a_0 = a, \ a_{n+1} = b.$$

If a set $a_1, a_2, \ldots, a_n$ is not finite the method of determining these numbers gives a sequence (a_n) which tends to a limit c where $c \leqq b$, because each element of the sequence is less than b and the sequence is monotonic and increasing. It will now be shown that the supposed sequence has no limiting point and that, in fact, c may be taken to be one of the numbers a_n.

The function $f(x)$ is continuous at c and therefore there is a number c_1 such that $|f(x) - f(c)| < \frac{1}{6}\varepsilon$ if $c_1 \leqq x \leqq c$. Again since, by hypothesis, c is a limiting point of the sequence (a_n) there is an element, a_m say, of the sequence such that $c_1 < a_m < c$ and therefore, by the last inequality, $|f(a_m) - f(c)| < \frac{1}{6}\varepsilon$. Hence if $a_m \leqq x \leqq c$

$$|f(x) - f(a_m)| \leqq |f(x) - f(c)| + |f(c) - f(a_m)| < \tfrac{1}{3}\varepsilon,$$

so that c may be taken to be the element a_{m+1}. The supposition therefore that the point b cannot be reached in a finite number of steps is untenable.

The interval (a, b) must be *closed*; if b were only a limiting point of the set of values of x and not itself a *value* of x the above reasoning would fail.

Suppose now that h is the least of the intervals (1), that is, that h is the least of the differences $(a_1 - a), (a_2 - a_1) \ldots, (b - a_n)$; then

$$|f(x') - f(x'')| < \varepsilon \text{ if } |x' - x''| < h.$$

For, *either* x' and x'' lie in the same interval, (a_r, a_{r+1}) say, and then

$$|f(x')-f(a_r)|<\tfrac{1}{3}\varepsilon,\ |f(x'')-f(a_r)|<\tfrac{1}{3}\varepsilon,$$

so that

$$|f(x')-f(x'')| \leqq |f(x')-f(a_r)| + |f(x'')-f(a_r)| < \tfrac{2}{3}\varepsilon < \varepsilon;$$

or else, x' and x'' lie in adjacent intervals (a_{r-1}, a_r) and (a_r, a_{r+1}). In this case, x' being in (a_{r-1}, a_r) and x'' in (a_r, a_{r+1}),

$$|f(x')-f(a_{r-1})|<\frac{\varepsilon}{3},\ |f(a_{r-1})-f(a_r)|<\frac{\varepsilon}{3},\ |f(x'')-f(a_r)|<\frac{\varepsilon}{3},$$

and therefore $\quad |f(x')-f(x'')|<\varepsilon.$

Thus $\quad |f(x')-f(x'')|<\varepsilon$ if $|x'-x''|<h.$

Uniform Continuity. This theorem expresses the property of *uniform* continuity. In virtue of the continuity of $f(x)$ it is possible to choose h_1 so that $|f(x)-f(c_1)|<\varepsilon$ if $|x-c_1|<h_1$, and also to choose h_2 so that $|f(x)-f(c_2)|<\varepsilon$ if $|x-c_2|<h_2$; but it is quite possible that h_2 would have to be less than h_1. The theorem however proves that, no matter what point c in (a, b) is taken, there is always *one* value of h such that $|f(x)-f(c)|<\varepsilon$ i $|x-c|<h$. The *uniformity* of the continuity lies in the fact that the *same* value of h secures the inequality $|f(x)-f(c)|<\varepsilon$ when $|x-c|<h$ *whatever point* in the interval (a, b) the point c may be.

THEOREM IV. *If $f(x)$ is continuous for the range $a \leqq x \leqq b$ it is bounded for that range.*

Let $a, a_1, a_2, \ldots, a_n, b$ be an increasing set of numbers that divide the interval (a, b) into $(n+1)$ sub-intervals such that $a_{r+1}-a_r<h$ and

$$|f(x)-f(a_r)|<\varepsilon \text{ if } |x-a_r|<h,\ r=0, 1, \ldots, n,\ a_0=a,\ a_{n+1}=b.$$

If $a_r<x\leqq a_{r+1}$ we have

$$f(x)=f(a)+\{f(a_1)-f(a)\}+\{f(a_2)-f(a_1)\}+\ldots+\{f(x)-f(a_r)\}$$

and therefore

$$|f(x)|\leqq|f(a)|+|f(a_1)-f(a)|+|f(a_2)-f(a_1)|+\ldots$$
$$+|f(x)-f(a_r)|<|f(a)|+(r+1)\varepsilon.$$

Now $(r+1)\varepsilon\leqq(n+1)\varepsilon$, a finite number, and therefore if $|f(a)|+(n+1)\varepsilon=k$ we have $|f(x)|<k$ when $a\leqq x\leqq b$ so that $f(x)$ is bounded.

The theorem follows at once from § 27. If $f(x)$ is not bounded in (a, b), there is at least one point ξ in the interval (a, b) in the neighbourhood of which $f(x)$ is not bounded; but this is impossible, because $f(x)$ is continuous at ξ and therefore $f(x)$ lies between $f(\xi)-\varepsilon$ and $f(\xi)+\varepsilon$, when x is any number in the interval $(\xi-\varepsilon, \xi+\varepsilon)$.

THEOREM V. *If $f(x)$ is continuous for the range $a \leqq x \leqq b$ then the upper bound M and the lower bound m of $f(x)$ are* values *of $f(x)$; or, $f(x)$* attains *its upper and lower bounds.*

By Theorem IV, $f(x)$ is bounded and therefore has an upper bound M and a lower bound m; it has to be proved that M and m are values that $f(x)$ actually has—that is, that there is at least one value ξ for which $f(\xi)=M$ and at least one value ξ' for which $f(\xi')=m$.

The Theorem of § 27 proves that there is at least one value ξ in the neighbourhood of which the upper bound of $f(x)$ is M. Now $f(x)$ is continuous at ξ and therefore, given ε as usual, there is a positive number h such that $|f(x)-f(\xi)|<\varepsilon$ if $|x-\xi|<h$. But M is the upper bound of $f(x)$ in the interval $(\xi-h, \xi+h)$ and therefore there is a value of x in this interval such that $M \geqq f(x) > M-\varepsilon$ or $M-f(x)<\varepsilon$. Hence

$$|M-f(\xi)|=|(M-f(x)+f(x)-f(\xi)| \leqq |M-f(x)|+|f(x)-f(\xi)|$$

so that $|M-f(\xi)|<2\varepsilon$. But M and $f(\xi)$ are constants and ε is arbitrarily small; therefore $M=f(\xi)$.

In the same way it is proved that $m=f(\xi')$ where $a \leqq \xi' \leqq b$.

M is *the maximum* and m *the minimum* value of $f(x)$.

29. Discontinuity. In § 44 of the *Elementary Treatise* the discontinuity of a function which is in general continuous is briefly referred to; Fig. 27, p. 88, and Fig. 32, p. 155, of that book are graphical representations of certain types of discontinuity. Fig. 32 should be specially considered, as it represents cases that actually occur and not cases manufactured to prove a possibility.

Removable Discontinuities. Suppose $f(x)$ to be defined for a range $a \leqq x \leqq b$; if $a<c<b$ it may happen that when $x \to c+0$ (that is, tends to c through values greater than c) $f(x)$ tends to a limit l and that when $x \to c-0$ (that is, tends to c through values less than c) $f(x)$ tends to the same limit l, but that l is

not the value $f(c)$ which the function has by its definition. The function is therefore discontinuous at c. In this case the definition of $f(x)$ for the value c of x may be changed and $f(c)$ taken to be equal to l; if c is the only point of discontinuity in an interval $(c-h,\ c+h)$ this change would make $f(x)$ continuous in the interval.

In this case the discontinuity is said to be *removable* and when the case occurs the change is usually made.

Discontinuities of the First Kind. If $f(x)$ tends to a limit l when $x \to c-0$ and also to a limit l' when x tends to $c+0$ and l is not equal to l', whether or not one of the numbers l, l' is equal to $f(c)$, the discontinuity is said to be of the *First Kind.* As a rule, $f(x)$ is not, by its original definition, defined for the value c of x, but in this case no value assigned to $f(c)$ will make $f(x)$ continuous at c. It is not unusual to define $f(c)$ to be $\frac{1}{2}(l+l')$—the mean of the two limits l and l' (*E.T.* Fig. 27, illustrates this type).*

Discontinuities of the Second Kind. If one (or both) of the limits of $f(x)$ when $x \to c-0$ and when $x \to c+0$ does not exist the discontinuity is said to be of the *Second Kind.* The function $\sin\left(\frac{1}{x-c}\right)$ illustrates this case; the function does not tend to a limit either when $x \to c-0$ or when $x \to c+0$.

30. Derivatives. If $f(x)$ is defined for the range $a \leqq x \leqq b$ and if x and x_1 are any two values of the argument in the range, $f(x)$ is said to have a derivative, denoted by $f'(x_1)$, for the value x_1 of the argument when the quotient $\varphi(x)$, where

$$\varphi(x) = \{f(x) - f(x_1)\}/(x - x_1), \quad \text{.....................}(1)$$

has a limit l for x tending to x_1. It is to be specially observed that the limit must be the same whether x tends to x_1+0 or to x_1-0. The number l is supposed to be *finite*; the cases $l = +\infty$ and $l = -\infty$ are considered a little further on.

When the number l exists $f(x)$ is said to be *differentiable at* x_1, or to have a derivative when $x = x_1$.

It may happen that the quotient $\varphi(x)$ tends to a limit l_1 when $x \to x_1+0$ and to a different limit l_2 when $x \to x_1-0$. In this case $f(x)$ is not differentiable at x_1 but $f(x)$ is said to have

* The letters "*E.T.*" indicate the *Elementary Treatise on the Calculus.*

at x_1 "a derivative on the right" or "a progressive derivative" l_1 and "a derivative on the left" or "a regressive derivative" l_2. Unless $l_1 = l_2$ the function $f(x)$ is not differentiable at x_1; if, however, x_1 is the extremity of an interval (a, b) the function $f(x)$ will be said to be differentiable in (a, b) if it is differentiable for every x between a and b and has a progressive derivative at a and a regressive at b.

For example, let $f(x)$ be defined as follows:

$$f(x) = 1 + x \text{ if } x \leqq 2, \text{ but } f(x) = 5 - x \text{ if } x \geqq 2.$$

Here $f(x)$ is differentiable for all values of x except for $x = 2$. The quotient $\{f(x) - f(2)\}/(x-2)$ tends to -1 when x tends to 2 from above but to $+1$ when x tends to 2 from below. There is a progressive derivative -1 and a regressive derivative $+1$ for the value 2 of x, but $f(x)$ is not differentiable for the value 2 of x.

If $f(x)$ does not tend to $f(x_1)$ when x tends to x_1 the limit l of the quotient $\varphi(x)$ does not exist and therefore $f(x)$ is not differentiable at x_1. Hence $f(x)$ *is not differentiable at* x_1 *unless* $f(x)$ *is continuous at* x_1; if $f(x)$ is differentiable for the range $a \leqq x \leqq b$ it must be continuous for that range.

The converse of this statement is, however, not true; that is, it is possible for $f(x)$ to be continuous for $a \leqq x \leqq b$ and yet not differentiable for any value of x in that range. See Hobson's *Functions of a Real Variable*, § 425 of First Edition. Non-differentiable functions of this character are outside our limits.

Cases $l = +\infty$ *and* $l = -\infty$. If $f(x) \to \infty$ when $x \to x_1$ the function is not continuous at x_1 and therefore has no derivative for $x = x_1$. On the other hand, if $f(x_1)$ is a finite number and if the quotient $\varphi(x)$ tends to $+\infty$ when x tends to x_1 (whether from above or from below) it is reasonable to say, especially in view of the geometrical interpretation of $f'(x_1)$ as a gradient, that $f(x)$ has a derivative, but that the derivative is $+\infty$. Similarly, if $\varphi(x)$ tends to $-\infty$ whether x tends to $x_1 - 0$ or to $x_1 + 0$, the derivative of $f(x)$ for $x = x_1$ is $-\infty$. *In all general theorems on derivatives, however, it is assumed that the limit* l *is finite*; each case of an infinite derivative must be considered by itself.

If $f(x) = (x - x_1)^{\frac{1}{3}}$ the derivative of $f(x)$ for $x = x_1$ is $+\infty$, but if $f(x) = (x - x_1)^{\frac{2}{3}}$ it has no derivative for $x = x_1$ since $\varphi(x)$ tends to $+\infty$ or to $-\infty$ according as x tends to $x_1 + 0$ or to $x_1 - 0$.

The student might with advantage read pp. 104-108 of the *Elementary Treatise* where various considerations respecting the derivative are stated. It is useful to remember that $f(x)$ is *strictly monotonic* for the range $a \leqq x \leqq b$ if for every value of x in that range $f(x)$ is continuous and has a derivative $f'(x)$ that is either always positive or else always negative when $a < x < b$. At a (or b) the derivative may be 0 or $+\infty$ or $-\infty$.

The Theorem of § 34, Ex. 4, should be noted.

Ex. 1. If $f(x) = x \sin \frac{1}{x}$ and if $f(x)$ is assigned the value 0 when $x=0$, show that $f(x)$ is continuous for $x=0$ but has no derivative for $x=0$.

It is necessary to assign a value to $f(x)$ when $x=0$ because $\sin(1/x)$ is undefined for $x=0$. The derivative of $f(x)$ for $x=0$, if it existed, would be the limit for x tending to 0 of $f(x)/x$, that is, of $\sin(1/x)$, so that there is no derivative for $x=0$.

Ex. 2. If $f(x) = x^2 \sin \frac{1}{x}$ and $f(0)=0$ show that $f(x)$ is differentiable for all values of x, on the understanding that $x \sin(1/x)$ is 0 when $x=0$.

Here $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ if x is not zero, but $f'(x)=0$ when $x=0$ since

$$f'(0) = \mathop{\mathcal{L}}_{x \to 0} \frac{f(x)}{x} = \mathop{\mathcal{L}}_{x \to 0} x \sin \frac{1}{x} = 0.$$

The derivative $f'(x)$ is discontinuous and has a discontinuity of the second kind at $x=0$; for we have

$$\mathop{\mathcal{L}}_{x \to 0} f'(x) = \mathop{\mathcal{L}}_{x \to 0} \left(2x \sin \frac{1}{x} - \cos \frac{1}{x}\right) = -\mathop{\mathcal{L}}_{x \to 0} \cos \frac{1}{x},$$

and $\cos(1/x)$ does not tend to a limit when $x \to 0$.

Ex. 3. If $f(x) = x \tanh \frac{1}{x}$ and $f(0)=0$ show that $f(x)$ is not differentiable for $x=0$, but has both a progressive and a regressive derivative for $x=0$.

31. Elementary Functions. Function of a Function. The derivatives of x^n, when n is any real number, and of e^x and $\log x$ have been considered in § 24; x in the case of e^x may be any real number while it may be any positive real number in the cases of x^n and $\log x$.

It is possible to define $\sin x$ and $\cos x$ by infinite series without any assumption of the geometrical meaning of the functions and when x is complex the functions are in fact defined by series (§ 69). It does not, however, seem to be

desirable to depart at this stage from the usual definitions, and it is not therefore necessary to reconsider the proofs of the derivatives of the direct trigonometric functions as these appear in the *Elementary Treatise*.

The theorems on *Function of a Function* and *Inverse Functions* may be noted; the first of these will be considered in this article and the second in the following article.

Function of a Function. If $y=f(x)$ and $x=\varphi(u)$, wher $\varphi(u)$ is single-valued and continuous for a given range of and $f(x)$ single-valued and continuous for the correspondin range of x, then y is said to be *a function of a function of u*. If $F(u)$ denote this function of u, that is, $F(u)=f[\varphi(u)]$, it will first be proved that $F(u)$ is a *continuous* function of u.

The function y or $f(x)$ is continuous and therefore, ε having the usual meaning, there is a positive number h such that

$$|y_1-y|=|f(x_1)-f(x)|<\varepsilon \text{ if } |x_1-x|<h.$$

Again, $\varphi(u)$ being continuous, there is a positive number k such that

$$|x_1-x|=|\varphi(u_1)-\varphi(u)|<h \text{ if } |u_1-u|<k.$$

Hence $\quad |F(u_1)-F(u)|=|y_1-y|<\varepsilon$ if $|u_1-u|<k$,

and therefore $F(u)$ is a continuous function of u.

Next suppose that the derivatives $f'(x)$ and $\varphi'(u)$ exist, and let x_1 and y_1 be the values of x and y corresponding to the value u_1 of u so that $x_1-x=\delta x=\delta\varphi(u)$ and $y_1-y=\delta y=\delta F(u)$. Two cases have to be considered.

(1) If δx is not zero for any value of u_1 in the neighbourhood of u we have the identity

$$\frac{\delta F(u)}{\delta u}=\frac{\delta f(x)}{\delta x}\frac{\delta\varphi(u)}{\delta u}, \quad\text{.......................}(a)$$

and therefore, since $f'(x)$ and $\varphi'(u)$ exist,

$$F'(u)=f'(x)\varphi'(u).$$

(2) Since x is a function of u and not an independent variable, it is possible that for one or more values of u_1 the increment δx may be zero and for such values of x the identity (a) would not be valid. But, by the definition of $f'(x)$,

$$\frac{\delta f(x)}{\delta x}=f'(x)+\alpha,$$

where $\alpha \to 0$ when $\delta x \to 0$, and therefore

$$\delta F(u) = \delta y = [f'(x) + \alpha]\,\delta x = [f'(x) + \alpha]\,\delta\varphi(u), \quad \ldots\ldots(a')$$

so that

$$\frac{\delta F(u)}{\delta u} = [f'(x) + \alpha]\frac{\delta\varphi(u)}{\delta u}.$$

When $\delta u \to 0$ so does α and the derivative $\varphi'(u)$ exists so that, letting δu tend to zero, we get the same value of $F'(u)$ as before.

It should be noted that equation (a') is true even if $\delta x = 0$ because δy is then also zero and the derivative $f'(x)$ exists. If δx is zero for an infinite number of values of u_1 in the neighbourhood of u so is $\delta x/\delta u$, that is $\delta\varphi(u)/\delta u$, and therefore since $\varphi'(u)$ exists $\varphi'(u)$ is zero. In this case $F'(u)$ is also zero.

32. Inverse Functions. Let $f(x)$ be a continuous, strictly monotonic function of x for the range $a \leqq x \leqq b$, that is, $f(x_2) > f(x_1)$ when $x_2 > x_1$ or else $f(x_2) < f(x_1)$ when $x_2 > x_1$; then, by Theorem II of § 28, $f(x)$ takes every value between $f(a)$ and $f(b)$ as x varies continuously from a to b, and can take each value only once since $f(x_1)$ and $f(x_2)$ are unequal when x_1 and x_2 are unequal. Hence the equation $f(x) = y$, where y lies between $f(a)$ and $f(b)$, has one and only one solution, say $x = \varphi(y)$, and therefore $\varphi(y)$ is a single-valued function of y. The function φ is called the *inverse* of the function f, and the equations $f[\varphi(y)] = y$ and $\varphi[f(x)] = x$ are identities (*E.T.* p. 18).

If $f(a) = a'$ and $f(b) = b'$ the function $\varphi(y)$ either *steadily* increases [that is, $\varphi(y_2) > \varphi(y_1)$ if $y_2 > y_1$] or else *steadily* decreases [that is, $\varphi(y_2) < \varphi(y_1)$ if $y_2 > y_1$] as y varies from a' to b'. Further, $\varphi(y)$ is *continuous.*

For, if $y_1 = f(x_1)$ and if x lies in the interval $(x_1 - h,\ x_1 + h)$, y will, since $f(x)$ is continuous, lie in an interval $(y_1 - \lambda',\ y_1 + \lambda'')$, and therefore if λ is the smaller of the two positive numbers λ' and λ'' the difference $|\,x - x_1\,|$ will be less than h when $|\,y - y_1\,|$ is less than λ. Hence $x \to x_1$, that is $\varphi(y) \to \varphi(y_1)$, when $y \to y_1$, and therefore $\varphi(y)$ is continuous.

Now let y and y' be two unequal numbers in the (closed) interval $(a',\ b')$ and let x and x' be the corresponding values of x, which are necessarily unequal; we now have the identity

$$\frac{x' - x}{y' - y} = 1 \div \left(\frac{y' - y}{x' - x}\right).$$

If $(y'-y)/(x'-x)$ tends to a limit that is not zero when x' tends to x, that is, if the derivative $f'(x)$ exists and is not zero, we deduce at once, since $x' \to x$ when $y' \to y$, that the derivative dx/dy or $\varphi'(y)$ exists and is given by the equation

$$\varphi'(y) = 1/f'(x).$$

If $f(x) > f(a)$ when $x > a$ the derivatives $f'(x)$ and $\varphi'(y)$ are both positive; hence if $f'(x) \to 0$ when $x \to a$ the derivative $\varphi'(y)$ will tend to $+\infty$ when y tends to a'. Similar considerations apply if $f'(x)$ is negative when $x > a$ and tends to zero when $x \to a$, and the cases in which $x \to b$ can be treated in like manner.

The derivatives of the inverse trigonometric functions may be found as in the *Elementary Treatise*, § 64. There is one change, however, that seems to be desirable, namely, that the range of $\cot^{-1} x$ should be from 0 to π and not from $-\dfrac{\pi}{2}$ to $\dfrac{\pi}{2}$; with the new convention

$$\tan^{-1} x + \cot^{-1} x = \frac{\pi}{2}.$$

33. Rolle's Theorem. A proof of this theorem will now be given in which the proposition (tacitly assumed *E.T.* p. 162) that a continuous function reaches, under certain conditions, its upper and lower bounds becomes one of the essential elements. The theorem may now be stated as follows:

If $F(x)$ is continuous in the closed interval (a, b) and has a derivative $F'(x)$ for the range $a < x < b$, that is, for the open interval (a, b); if further $F(a) = 0$ and $F(b) = 0$, then $F'(x)$ will be zero for at least one value ξ where $a < \xi < b$.

Of course $F(x)$ is continuous for those values for which $F'(x)$ exists, but for the validity of the proof it is necessary that, when x tends to a or to b from within the interval, $F(x)$ should tend to zero; the particular form given to the enunciation of the theorem secures this.

If $F(x)$ is constantly zero $F'(x)$ is also zero for $a < x < b$. If $F(x)$ is not constantly zero it must take either positive or negative values or both, and therefore, *being a continuous function*, must actually take for at least one value of x its upper bound, if $F(x)$ is positive, and its lower bound if $F(x)$ is negative. Suppose that $F(x)$ takes positive values; then, for at least

one value ξ such that $a<\xi<b$, $F(x)$ is equal to M, its upper bound. If the positive number h is sufficiently small both $F(\xi+h)$ and $F(\xi-h)$ will be less than $F(\xi)$, and of the two quotients

$$\frac{F(\xi+h)-F(\xi)}{h} \text{ and } \frac{F(\xi-h)-F(\xi)}{-h},$$

the first will be negative and the second positive. Now the derivative $F'(\xi)$ exists and is the limit for h tending to zero of either quotient; as the limit of the first quotient $F'(\xi)$ is negative if not zero, while as the limit of the second it is positive if not zero. The only possible conclusion is therefore that $F'(\xi)$ is zero, as was to be proved. The same conclusion follows if $F(x)$ takes negative values, since it must be equal to its lower bound for at least one x such that $a<x<b$.

The proof does not require that $F'(x)$ should be *finite*, only that it should be *definite*; *geometrically*, the graph of $F(x)$ might have an inflexional tangent at $[\xi, F(\xi)]$ perpendicular to the x-axis $(a<\xi<b)$. But $F(x)$ must be continuous.

The following method of discussing the theorem depends on the use of the derivative as a test of an increasing or decreasing function.

By the definition of the derivative

$$F(x+h)-F(x)=h\{F'(x)+\lambda\}$$

where $\lambda \rightarrow 0$ when $h \rightarrow 0$ so that if $|h|$ is sufficiently small the sign of $F'(x)+\lambda$ is that of $F'(x)$ provided $F'(x)$ is not zero. Hence, so long as $F'(x)$ is positive $F(x)$ increases or decreases according as x increases or decreases, while so long as $F'(x)$ is negative $F(x)$ decreases or increases according as x increases or decreases. Conversely, $F'(x)$ not being zero, if, for example, $F(x)$ increases as x increases by $|h|$ the derivative $F'(x)$ must be positive, but if $F(x)$ decreases when x increases by $|h|$ the derivative $F'(x)$ must be negative, $|h|$ being sufficiently small.

Now if $F(x)$ takes positive values in the interval (a, b) it must, since $F(a)=0$, $F(b)=0$ and $F(x)$ is continuous, have an upper bound M which it reaches for a value ξ of x between a and b. Hence $F'(x)$ is positive if $\xi-h<x<\xi$ and negative if $\xi<x<\xi+h$ when the positive number h is sufficiently small. If $F'(\xi)$ is not zero $F(x)$ will either increase from a value a

little less than $F(\xi)$ to one a little greater than $F(\xi)$, or else decrease from one that is greater than $F(\xi)$ to one that is less as x increases from $\xi - h$ to $\xi + h$. But there is no greater value than $F(\xi)$ and therefore $F'(\xi) = 0$.

The same argument holds if $F(x)$ takes negative values in the interval (a, b).

Ex. 1. Suppose that $f(x)$ is continuous and has a derivative $f'(x)$ for the range $a \leqq x \leqq b$. If $f'(a)$ and $f'(b)$ are unequal and if k is any number between $f'(a)$ and $f'(b)$ there is a value ξ such that $f'(\xi) = k$ where $a < \xi < b$.

Let $\varphi(x) = f(x) - kx$; then $\varphi(x)$ is continuous and has a derivative $\varphi'(x)$, equal to $f'(x) - k$. Now $\varphi'(a)$ and $\varphi'(b)$ have opposite signs since k lies between $f'(a)$ and $f'(b)$. Suppose $\varphi'(a) > 0$ and $\varphi'(b) < 0$. Since $\varphi'(a)$ is positive, $\varphi(x)$ increases as x increases from a, and since $\varphi'(b)$ is negative $\varphi(x)$ also increases as x decreases from b. Now $\varphi(x)$ is continuous and therefore has an upper bound G which it attains for a value, ξ say, between a and b; but if $\varphi(\xi)$ is the upper bound $\varphi'(\xi) = 0$ and therefore $f'(\xi) = k$. Similarly it is seen that if $\varphi'(a) < 0$ there is a lower bound for $\varphi(x)$ and therefore a value of x for which $\varphi'(x) = 0$ or $f'(x) = k$.

Ex. 2. If $F(x)$ and $F'(x)$ satisfy the conditions of Rolle's Theorem for the interval (a, b) and if α and β are two values of x in the interval such that $F(\alpha) = F(\beta)$ and $\alpha < \beta$, show that there is a value ξ such that $F'(\xi) = 0$ where $\alpha < \xi < \beta$.

Ex. 3. If $a_1 < a_2 < a_3 < \ldots < a_n$ and if $F(x)$ and its derivatives up to and including the $(n-1)$th derivative are continuous for the range $a_1 \leqq x \leqq a_n$ prove that when $F(a_1)$, $F(a_2)$, $F(a_3)\ldots$, $F(a_n)$ are each zero, $F^{(n-1)}(x)$ will vanish for at least one value of x in the interval (a_1, a_n).

34. Theorem of Mean Value. This theorem (*E.T.* pp. 162-165) is an immediate deduction from Rolle's Theorem and may be stated as follows: If $f(x)$ is continuous in the closed interval (a, b) and has a derivative $f'(x)$ for every value of x in the open interval (a, b), then

$$f(b) = f(a) + (b-a)f'(\xi), \qquad \text{(1)}$$

where $a < \xi < b$.

Take the function $F(x)$ so that

$$F(x) = f(x) - f(a) - \frac{x-a}{b-a}\{f(b) - f(a)\}.$$

$F(x)$ satisfies the conditions of Rolle's Theorem, and therefore there is at least one value ξ of x such that $F'(\xi) = 0$ and $a < \xi < b$; thus

$$f'(\xi) - \{f(b) - f(a)\}/(b-a) = 0,$$

that is,

$$f(b) = f(a) + (b-a)f'(\xi).$$

An immediate deduction from this equation is that if $f'(x)$ is zero when $a<x<b$ the function $f(x)$ is constant for that range; for if c and d are any two such values of x the theorem is applicable, and therefore, ξ being some number between c and d,

$$f(d)=f(c)+(d-c)f'(\xi)=f(c).$$

Thus all the values of $f(x)$ in question are the same.

It follows at once that if $f(x)$ and $\varphi(x)$ have derivatives that are equal for every value of x in (a, b) the functions differ, if at all, by a constant; for if $F(x)=f(x)-\varphi(x)$ the derivative $F'(x)$ is zero.

The equation (1) may be put in different forms (*E.T.* § 73); a useful form is

$$f(a+h)=f(a)+hf'(a+\theta h), \quad 0<\theta<1. \quad \text{............(2)}$$

The theorem in Ex. 4 should be noted.

Ex. 1. If the functions $f(x)$, $\varphi(x)$, $\psi(x)$ are defined for the closed interval (a, b) and have derivatives for the open interval (a, b) prove that

$$\begin{vmatrix} f(a) & \varphi(a) & \psi(a) \\ f(b) & \varphi(b) & \psi(b) \\ f'(\xi) & \varphi'(\xi) & \psi'(\xi) \end{vmatrix} = 0,$$

where $a<\xi<b$, and deduce the theorem of *E.T.* p. 419.

Let $F(x)$ be the determinant formed from the given determinant by putting $f(x)$, $\varphi(x)$, $\psi(x)$ in place of $f'(\xi)$, $\varphi'(\xi)$, $\psi'(\xi)$ respectively; $F(x)$ will satisfy the conditions of Rolle's Theorem and $F'(\xi)$ is the given determinant. Next let $f(x)=1$; then if $\psi'(x)$ is not zero for $a<x<b$ we find

$$\frac{\varphi(b)-\varphi(a)}{\psi(b)-\psi(a)}=\frac{\varphi'(\xi)}{\psi'(\xi)}.$$

Ex. 2. If $f(x)$ is continuous for the closed interval (a, b) and has a derivative $f'(x)$ which is *bounded* for the open interval (a, b), say $|f'(x)|<K$, then $|f(x_2)-f(x_1)|<K\,|x_2-x_1|$ where x_1 and x_2 are any two values of x within the interval (a, b).

The Theorem of Mean Value is applicable under the conditions required by Rolle's Theorem; these conditions do not require that $f'(x)$ should be finite but only that it should be a definite number, finite or $+\infty$ or $-\infty$. If, however, $f'(x)$ is bounded and if x_1, x_2 are any two numbers in (a, b) we have

$$f(x_2)=f(x_1)+(x_2-x_1)f'(\xi), \quad x_1<\xi<x_2 \text{ or } x_2<\xi<x_1,$$

and therefore

$$|f(x_2)-f(x_1)|=|(x_2-x_1)f'(\xi)|<K\,|x_2-x_1|$$

if $|f'(x)|<K$ when $a<x<b$.

A function $f(x)$ which satisfies the condition

$$|f(x_2)-f(x_1)|<K|x_2-x_1|$$

when x_1 and x_2 are any two values of x in (a, b) is sometimes said to satisfy "**Lipschitz's Condition.**"

Ex. 3. If $f(x)$ has a derivative $f'(x)$ in the interval (a, b) and if c is a point in (a, b) such that $f'(x)$ tends to l when x tends to c then $l=f'(c)$.

For,

$$\frac{f(c+h)-f(c)}{h}=f'(c+\theta h)\ ;\ \text{ but } f'(c+\theta h)\to l \text{ and } \frac{f(c+h)-f(c)}{h}\to f'(c).$$

Ex. 4. Theorem. *If* f′(x) *is a continuous function of* x *for the range* $\mathrm{a}\leqq \mathrm{x}\leqq \mathrm{b}$ *the quotient* $\{\mathrm{f}(\mathrm{x}+\mathrm{h})-\mathrm{f}(\mathrm{x})\}/\mathrm{h}$ *converges* uniformly *to* f′(x) *when* h *tends to zero.*

By the Mean Value Theorem

$$f(x+h)-f(x)=hf'(x+\theta h)$$

and therefore

$$\frac{f(x+h)-f(x)}{h}-f'(x)=f'(x+\theta h)-f'(x).$$

Now, since $f'(x)$ is continuous it is uniformly continuous (§ 28, Th. III), and therefore, given ϵ as usual, there is a positive number k such that

$$|f'(x+h)-f'(x)|<\epsilon, \text{ if } |h|<k$$

whatever value x may have in the interval. Hence, since $|\theta h|<|h|$,

$$\left|\frac{f(x+h)-f(x)}{h}-f'(x)\right|<\epsilon \text{ if } |h|<k,$$

so that the convergence is uniform—*i.e.* does not depend on x.

EXERCISES III.

1. The functions $f(x)$ and $f'(x)$ are continuous for the range $a\leqq x\leqq b$. If $f(a)$ and $f(b)$ are zero but $f'(a)$ and $f'(b)$ not zero prove that, whatever number k may be, $f(x)=kf'(x)$ for at least one value of x between a and b.

2. Prove that e^x-1 is greater than $(1+x)\log(1+x)$ if x is positive.

3. If $f(x)=e^x(x^2-6x+12)-(x^2+6x+12)$ find $f'(x)$, $f''(x)$, $f'''(x)$, and show that $f'(x)$ is positive when x is positive. Deduce that when x is positive

$$\frac{1}{e^x-1}-\frac{1}{x}+\frac{1}{2}<\frac{x}{12}.$$

Show also that the expression on the left of this inequality is positive when x is positive, and tends to zero when x tends to zero.

4. If $f(x)=e^x(x^2-x+2)-(x^2+x+2)$ prove that when x is positive $f(x)$ increases as x increases. Deduce that if

$$\varphi(x)=\frac{e^x(x-2)+(x+2)}{x^2(e^x-1)},$$

$\varphi(x)$ tends to $\frac{1}{6}$ when x tends to zero, and that $\varphi(x)$ cannot be greater than $\frac{1}{6}$ whatever value x may have. (Hermite.)

5. If $a < x < b$, prove that

$$f(x) = \frac{(x-a)f(b)+(b-x)f(a)}{b-a} - \tfrac{1}{2}(b-x)(x-a)f''(x_1),$$

where $a < x_1 < b$.

6. Show that if $\psi''(x)$ is not zero for $a < x < b$,

$$\frac{\varphi(b)-\varphi(a)-(b-a)\varphi'(a)}{\psi(b)-\psi(a)-(b-a)\,\psi'(a)} = \frac{\varphi''(x_1)}{\psi''(x_1)},\ a < x_1 < b.$$

7. Show that if $\psi^{(n)}(x)$ is not zero for $a < x < b$

$$\frac{f}{g} = \frac{\varphi^{(n)}(x_1)}{\psi^{(n)}(x_1)},\ a < x_1 < b,$$

where
$$f = \varphi(b) - \varphi(a) - \sum_{r=1}^{n-1} \frac{(b-a)^r}{r!}\varphi^{(r)}(a),$$

$$g = \psi(b) - \psi(a) - \sum_{r=1}^{n-1} \frac{(b-a)^r}{r!}\psi^{(r)}(a).$$

8. Determine the constants c_0, c_1, c_2 so that the quadratic function $Q(x)$ where

$$Q(x) = c_0 + c_1(x-a_1) + c_2(x-a_1)(x-a_2)$$

may be equal to $f(a_1), f(a_2), f(a_3)$ when x is equal to a_1, a_2, a_3 respectively, the numbers a_1, a_2, a_3 being all different ; then prove that

$$f(x) = Q(x) + \frac{f'''(x_1)}{3!}(x-a_1)(x-a_2)(x-a_3),$$

where x_1 lies between the least and the greatest of a_1, a_2, a_3, x.

[$c_0 = f(a_1)$, $\quad c_1 = \dfrac{f(a_1)}{a_1-a_2} + \dfrac{f(a_2)}{a_2-a_1}$,

$$c_2 = \frac{f(a_1)}{(a_1-a_2)(a_1-a_3)} + \frac{f(a_2)}{(a_2-a_1)(a_2-a_3)} + \frac{f(a_3)}{(a_3-a_1)(a_3-a_2)}.$$

Next choose P so that

$$f(x) = Q(x) + P(x-a_1)(x-a_2)(x-a_3),$$

and let
$$F(z) = f(z) - [Q(z) + P(z-a_1)(z-a_2)(z-a_3)].$$

The function $F(z)$ is zero for the values a_1, a_2, a_3, x of z, and therefore $F'''(z)$ vanishes for a value x_1 of z between the least and the greatest of a_1, a_2, a_3, x. But

$$F'''(z) = f'''(z) - 1\,.\,2\,.\,3P,$$

so that $P = f'''(x_1)/3!$]

9. Determine the constants $c_0, c_1, \ldots c_{n-1}$ so that the polynomial $Q(x)$ where

$$Q(x) = c_0 + c_1(x-a_1) + c_2(x-a_1)(x-a_2) + \ldots$$
$$+ c_{n-1}(x-a_1)(x-a_2)\ldots(x-a_{n-1})$$

shall be equal to $f(a_1), f(a_2) \ldots, f(a_n)$ when x is equal to $a_1, a_2, \ldots, a_n$ respectively, the numbers $a_1, a_2, \ldots, a_n$ being all different; then prove that

$$f(x) = Q(x) + \frac{f^{(n)}(x_1)}{n!}(x-a_1)(x-a_2)\ldots(x-a_n),$$

where x_1 lies between the least and the greatest of $a_1, a_2, \ldots, a_n, x$.

[$c_0 = f(a_1)$; c_r depends only on $a_1, a_2, \ldots a_{r+1}$ and if

$$\varphi_{r+1}(x) = (x-a_1)(x-a_2)\ldots(x-a_{r+1}),$$

$$c_r = \sum_{s=1}^{r+1} \frac{f(a_s)}{\varphi'_{r+1}(a_s)}.$$

The following notation is often used:

$$c_0 = f(a_1),\ c_1 = f(a_1, a_2),\ c_2 = f(a_1, a_2, a_3) \ldots c_r = f(a_1, a_2, \ldots a_r, a_{r+1}) \ldots,$$

and it may be proved that

$$f(a_1, a_2, \ldots a_r, a_{r+1}) = \frac{f(a_1, a_2, \ldots a_{r-1}, a_r) - f(a_1, a_2, \ldots a_{r-1}, a_{r+1})}{a_r - a_{r+1}}\Big].$$

10. If

$$f(x+h) - f(x) = \Delta f(x)$$
$$\Delta f(x+h) - \Delta f(x) = \Delta^2 f(x)$$

$$\Delta^{n-1} f(x+h) - \Delta^{n-1} f(x) = \Delta^n f(x)$$

and if, in Example 9, $a_1 = a$, $a_{r+1} = a + rh$, $r = 1, 2, \ldots,$ prove that $f(x)$ is equal to

$$f(a) + \frac{x-a}{h}\Delta f(a) + \frac{(x-a)(x-a-h)}{1\,.\,2\,h^2}\Delta^2 f(a) + \ldots$$
$$+ \frac{(x-a)(x-a-h)\ldots[x-a-(n-2)h]}{(n-1)!\,h^{n-1}}\Delta^{n-1} f(a)$$
$$+ \frac{(x-a)(x-a-h)\ldots[x-a-(n-1)h]}{n!\,h^n} f^{(n)}(x_1),$$

where x_1 lies between the least and the greatest of the numbers a, $a + (n-1)h$, x.

11. The equation $e^{x/k} - x = 0$ has no real roots if $k < e$, and never has more than 2 real roots for any real value of k.

12. If $0 < \alpha < \pi/2$ and $0 < x < \pi$, the equation

$$\sin(x - \alpha) = m \sin^4 x$$

where m is positive, has (i) one real root if $\tan \alpha > \frac{4}{3}$, and (ii) one or three real roots if $\tan \alpha < \frac{4}{3}$. There are three real roots if m lies between the minimum and the maximum values of the function $\sin(x-\alpha)/\sin^4 x$.

(Tisserand.)

13. If
$$y=\tan^{-1}x\cdot\frac{13x^3+3x}{3x^4+14x^2+3},$$
show that
$$\frac{dy}{dx}=\frac{16x^4(3x^2+1)(x^2-1)}{(1+x^2)(3x^4+14x^2+3)^2},$$
and then prove that the equation $y=0$ has three real roots. (Tisserand.)

35. Differentials. If $y=f(x)$ the derivative $f'(x)$ is the limit for $\delta x\to 0$ of $\delta y/\delta x$ so that

$$\frac{\delta y}{\delta x}=f'(x)+\lambda;\quad \delta y=f'(x)\delta x+\lambda\delta x,$$

where $\lambda\to 0$ when $\delta x\to 0$.

When x is the independent variable the part $f'(x)\delta x$ of δy is called the *differential* of y or $f(x)$, and is denoted by dy or $df(x)$. If δx is an infinitesimal (*E.T.* p. 195) the difference $(\delta y-dy)$ is an infinitesimal of a higher order, because $(\delta y-dy)/\delta x$ is equal to λ and $\lambda\to 0$ when $\delta x\to 0$. When in any calculation powers of δx higher than the first are to be rejected dy may be substituted for δy.

If x is a function $\varphi(t)$ of t then y is a function of t, say $y=f[\varphi(t)]=F(t)$; the independent variable is now t and therefore
$$dy=F'(t)\delta t=f'(x)\varphi'(t)\delta t.$$
But x is now a function of t and $dx=\varphi'(t)\delta t$, so that
$$dy=f'(x)\varphi'(t)\delta t=f'(x)dx.$$

Thus, when x is the independent variable $dy=f'(x)\delta x$, but when t is the independent variable $dy=f'(x)dx$. The two expressions for dy will therefore have the same form, whether x is the independent variable or not, provided we take δx to mean the same thing as dx when x is the independent variable. There can be no objection to doing so, since δx may be any number whatever *provided* $dy/\delta x$ is equal to $f'(x)$; but, further, δx and dx are the same thing when the function $f(x)$ is x itself because in that case $f'(x)=1$ and therefore $df(x)=dx$.

No confusion therefore can arise if the increment δx of the *independent variable* be denoted by dx. There is, besides, the notable advantage that dy has now the same form,
$$dy=f'(x)dx,\ \dots\dots\dots\dots\dots\dots(1)$$
whether x be the independent variable or not; it is this property of the differential that makes it so useful.

Thus the differential $d(uv)$ of the product uv is given by the equation
$$d(uv) = vdu + udv,$$
because, if u and v are functions of x,
$$\frac{d(uv)}{dx} = v\frac{du}{dx} + u\frac{dv}{dx}$$
and
$$d(uv) = v\frac{du}{dx}dx + u\frac{dv}{dx}dx = vdu + udv,$$
since $du = \frac{du}{dx}dx$ and $dv = \frac{dv}{dx}dx$, whether x be independent or not.

If $n>1$ the differential of the nth order $d^n y$, when y is a function of x, is defined by the equation
$$d^n y = f^{(n)}(x)(dx)^n = f^{(n)}(x)dx^n, \quad\text{.................}(2)$$
when x is the independent variable. If x be a function of t, say $x = \varphi(t)$ so that $y = f[\varphi(t)] = F(t)$, then
$$d^2y = F''(t)dt^2 \text{ and } d^2x = \varphi''(t)\,dt^2.$$
Now
$$F''(t) = \frac{d^2f(x)}{dx^2}\left[\frac{d\varphi(t)}{dt}\right]^2 + \frac{df(x)}{dx}\frac{d^2\varphi(t)}{dt^2},$$
so that
$$F''(t)dt^2 = \frac{d^2f(x)}{dx^2}\left[\frac{d\varphi(t)}{dt}dt\right]^2 + \frac{df(x)}{dx}\left[\frac{d^2\varphi(t)}{dt^2}dt^2\right]$$
$$= \frac{d^2f(x)}{dx^2}\,.\,dx^2 + \frac{df(x)}{dx}d^2x,$$
or
$$d^2y = f''(x)dx^2 + f'(x)d^2x, \quad\text{...........................}(3)$$
and this is different from the form $f''(x)dx^2$ which is the value of d^2y when x is the independent variable *unless* $d^2x = 0$. There is no longer the advantage of the same form for d^2y whether the variable x is or is not the independent variable, and the definition (2) is essentially confined to the case in which x is the independent variable, or, what is equivalent, to the assumption that dx is constant so that the differential of dx, that is, $d(dx)$ or d^2x, and all higher differentials of x are zero.

When $x = \varphi(t)$ and t is the independent variable
$$dx^n = [\varphi'(t)]^n dt^n$$
so that
$$d(dx^n) = n[\varphi'(t)]^{n-1}\varphi''(t)\,dt\,.\,dt^n$$
$$= n[\varphi'(t)\,dt]^{n-1}\varphi''(t)\,dt^2$$
and therefore
$$d(dx^n) = n\,dx^{n-1}\,d^2x.$$

If x is not independent we find by taking the differential of the product $f'(x)dx$ that

$$\begin{aligned} d[f'(x)dx] &= dxd[f'(x)] + f'(x)d(dx) \\ &= dxf''(x)dx + f'(x)d^2x \\ &= f''(x)dx^2 + f'(x)d^2x, \end{aligned}$$

so that the value of d^2y in (3) may be found by taking the differential of dy, that is, of $f'(x)dx$; if x is the independent variable dx is constant and $d^2x=0$.

In the same way

$$\begin{aligned} d^3y = d(d^2y) &= dx^2d[f''(x)] + f''(x)d[dx^2] + d^2xd[f'(x)] \\ &\qquad + f'(x)d[d^2x] \\ &= f'''(x)dx^3 + 3f''(x)dxd^2x + f'(x)d^3x. \end{aligned}$$

Similarly d^4y, d^5y, ... may be found.

Ex. A curve is given by the equations

$$x=f(t), \quad y=g(t), \quad z=h(t);$$

find the equations of the tangent at the point $P(x, y, z)$ and the equation of the plane to which the plane through the tangent at P and a point Q on the curve tends, as its limiting position, when Q tends along the curve to P.

The direction cosines of the chord through P, "the point t," and P' "the point $t+\delta t$," are proportional to

$$f(t+\delta t)-f(t), \quad g(t+\delta t)-g(t), \quad h(t+\delta t)-h(t),$$

that is, to $$f'(t)+\lambda_1, \quad g'(t)+\lambda_2, \quad h'(t)+\lambda_3$$

where λ_1, λ_2, λ_3 tend to zero when δt tends to zero. Hence, if ξ, η, ζ are current coordinates the equations of the tangent at P are

$$\frac{\xi-x}{f'(t)}=\frac{\eta-y}{g'(t)}=\frac{\zeta-z}{h'(t)},$$

or, if differentials be used,

$$\frac{\xi-x}{dx}=\frac{\eta-y}{dy}=\frac{\zeta-z}{dz}. \qquad \text{(i)}$$

The equation of a plane through the tangent at P is of the form

$$A(\xi-x)+B(\eta-y)+C(\zeta-z)=0. \qquad \text{(ii)}$$

where $$Adx+Bdy+Cdz=0. \qquad \text{(iii)}$$

If $t+dt$ is the parameter of the point Q these two equations must be satisfied when for x, y, z we put $x+\delta x$, $y+\delta y$, $z+\delta z$, and since the limit for $dt \to 0$ is alone required we may simply take the differential of each equation; therefore

$$-Adx-Bdy-Cdz=0$$

$$Ad^2x+Bd^2y+Cd^2z=0. \qquad \text{(iv)}$$

The first of these equations is merely (iii); elimination of A, B, C between (ii), (iii) and (iv) gives the required solution

$$\begin{vmatrix} \xi - x & \eta - y & \zeta - z \\ dx & dy & dz \\ d^2x & d^2y & d^2z \end{vmatrix} = 0.$$

36. Higher Derivatives. An expression for the nth derivative of a function of x is, when the function is at all complicated, usually difficult to find, though the value for $x=0$ may sometimes be easily obtained (*E.T.* p. 397). Some general methods have been elaborated and a statement of some of these will now be given; for further information the student may consult the books named below * from which the following exposition is largely drawn.

The nth derivative of y when $y=f(u)$ and $u=\varphi(x)$ is usually to be found by calculating a few successive derivatives; by noting the form it may be observed whether any law is suggested, the suggestion being then tested by mathematical induction. In the present case we have

$$\frac{dy}{dx}=\varphi'(x)f'(u), \quad \frac{d^2y}{dx^2}=\varphi''(x)f'(u)+[\varphi'(x)]^2f''(u);$$

and so on. It is at once suggested that the nth derivative will be an expression of the form

$$\frac{d^ny}{dx^n}=A_{n,1}f'(u)+\frac{A_{n,2}}{2!}f''(u)+\dots+\frac{A_{n,n}}{n!}f^{(n)}(u)$$

$$=\sum_{r=1}^{n}\frac{A_{n,r}}{r!}f^{(r)}(u) \quad \dots\dots\dots(1)$$

where the coefficients $A_{n,r}$ do not depend on the function $f(u)$ and will therefore be the same whatever function $f(u)$ may be, so long as $\varphi(x)$ is the same.

Now put for $f(u)$ successively $u, u^2, u^3, \dots, u^n$ in the equation (1); then

$$\frac{d^n\,.\,u}{dx^n}=A_{n,1}; \quad \frac{d^n\,.\,u^2}{dx^n}=A_{n,1}\,.\,2u+A_{n,2};$$

$$\frac{d^n\,.\,u^3}{dx^n}=A_{n,1}\,.\,3u^2+A_{n,2}\,.\,3u+A_{n,3}; \text{ and so on.}$$

* Schlömilch, *Compendium der höheren Analysis*, vol. 2, and *Übungsbuch zum Studium der höheren Analysis*, vol. 1; Nielsen, *Elemente der Funktionentheorie*; Tisserand, *Recueil complémentaire d'Exercices sur le Calcul Infinitésimal*.

These equations give $A_{n,1}$, then $A_{n,2}$, then $A_{n,3}$, and finally $A_{n,n}$, and the values obtained suggest the law

$$A_{n,r}=\sum_{s=0}^{r-1}(-1)^s\binom{r}{s}u^s D_x^n(u^{r-s}), \quad\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

where the symbol $\binom{r}{s}$ is the usual binomial coefficient ${}_rC_s$,

that is, $\qquad r(r-1)(r-2)\ldots(r-s+1)/s!$

To test (2), put $n+1$ for n and (2) becomes

$$A_{n+1,r}=\sum_{s=0}^{r-1}(-1)^s\binom{r}{s}u^s D_x^{n+1}(u^{r-s}). \quad\ldots\ldots\ldots\ldots\ldots(2')$$

Now differentiate (1); then

$$\frac{d^{n+1}y}{dx^{n+1}}=\sum_{r=1}^{n+1}\left\{\frac{1}{r!}\frac{dA_{n,r}}{dx}+\frac{A_{n,r-1}}{(r-1)!}\frac{du}{dx}\right\}f^{(r)}(u). \quad\ldots\ldots(3)$$

on the understanding that $A_{n,0}$ and $A_{n,n+1}$ are identically zero.

If we have the relation

$$\frac{dA_{n,r}}{dx}+rA_{n,r-1}\frac{du}{dx}=A_{n+1,r}, \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

equation (3) will show the same law as equation (1), and therefore the formula (2) will hold for every value of n. Now

$$\frac{dA_{n,r}}{dx}=\sum_{s=0}^{r-1}(-1)^s\binom{r}{s}u^s D_x^{n+1}(u^{r-s})$$
$$+\left\{\sum_{s=1}^{r-1}(-1)^s\binom{r}{s}su^{s-1}D_x^n(u^{r-s})\right\}\frac{du}{dx};$$

and the expression within the brackets is easily found to be

$$-r\sum_{s=0}^{r-2}(-1)^s\binom{r-1}{s}u^s D_x^n(u^{r-1-s})=-rA_{n,r-1},$$

because $\binom{r}{s}s=r\binom{r-1}{s-1}$ and the variable index s may be changed to $s+1$. Hence

$$\frac{dA_{n,r}}{dx}+rA_{n,r-1}\frac{du}{dx}=\sum_{s=0}^{r-1}(-1)^s\binom{r}{s}u^s D_x^{n+1}(u^{r-s})=A_{n+1,r},$$

and the relation (4) is established; the formula (1) where $A_{n,r}$ is given by (2) is therefore proved.

The expression for $A_{n,r}$ may be put in another form which is frequently more convenient. Let it be first noted that if $\varphi(x+\varrho)$ is a function of the sum $(x+\varrho)$ of the independent variables x and ϱ, we have

$$\frac{\partial^n\varphi(x+\varrho)}{\partial x^n}=\frac{\partial^n\varphi(x+\varrho)}{\partial \varrho^n}, \text{ so that } \frac{d^n\varphi(x)}{dx^n}=\left[\frac{\partial^n\varphi(x+\varrho)}{\partial\varrho^n}\right]_{\rho=0}.$$

Now let $[\varphi(x+\varrho)-\varphi(x)]^r$ be expanded by the binomial theorem; we get

$$\frac{\partial^n}{\partial\varrho^n}\left\{\varphi(x+\varrho)-\varphi(x)\right\}^r=\frac{\partial^n}{\partial\varrho^n}\sum_{s=0}^{r}(-1)^s\binom{r}{s}[\varphi(x)]^s[\varphi(x+\varrho)]^{r-s}$$

$$=\sum_{s=0}^{r-1}(-1)^s\binom{r}{s}[\varphi(x)]^s\frac{\partial^n\,.\,[\varphi(x+\varrho]^{r-s}}{\partial\varrho^n},$$

where the term for $s=r$ disappears since it is independent of ϱ.

But $$\left[\frac{\partial^n\,.\,[\varphi(x+\varrho)]^{r-s}}{\partial\varrho^n}\right]_{\rho=0}=\frac{d^n\,.\,[\varphi(x)]^{r-s}}{dx^n}=\frac{d^n\,.\,u^{r-s}}{dx^n}$$

and therefore

$$A_{n,r}=\left[\frac{\partial^n}{\partial\varrho^n}\left\{\varphi(x+\varrho)-\varphi(x)\right\}^r\right]_{\rho=0}.\qquad(5)$$

Ex 1. If $y=f(u)$ and $u=\varphi(x)=x^2$ find $\dfrac{d^ny}{dx^n}$.

In the formula (1) put $n-r$ for r; this choice of the variable of summation often gives a simpler form to the result; then

$$\frac{d^ny}{dx^n}=\sum_{r=0}^{n-1}\frac{A_{n,n-r}}{(n-r)!}f^{(n-r)}(u).$$

Now $$\{\varphi(x+\varrho)-\varphi(x)\}^{n-r}=(2x\varrho+\varrho^2)^{n-r};$$

the only term in the expansion which does not vanish, when ϱ is made zero after the differentiation with respect to ϱ, is that whose index is n and the nth derivative of ϱ^n is $n!$. Thus we find

$$\frac{A_{n,n-r}}{(n-r)!}=\binom{n-r}{r}\frac{n!}{(n-r)!}(2x)^{n-2r},$$

$$=\frac{n(n-1)\dots(n-2r+1)}{1\,.\,2\,.\dots r}(2x)^n$$

on the understanding that this expression is $(2x)^n$ when $r=0$; and therefore when $y=f(u)=f(x^2)$

$$\frac{d^ny}{dx^n}=\sum_{r=0}^{m}\frac{n(n-1)\dots(n-2r+1)}{1\,.\,2\dots r}(2x)^{n-2r}f^{(n-r)}(u),\qquad(6)$$

where m is $\frac{1}{2}n$ or $\frac{1}{2}(n-1)$ according as n is even or odd.

If $f(u)=(1-u)^{n-\frac{1}{2}}$ and if in (6) we put $n-1$ for n we find

$$\frac{d^{n-1}.(1-x^2)^{n-\frac{1}{2}}}{dx^{n-1}}$$
$$=\frac{(-1)^{n-1}K}{n}\sum_{r=0}(-1)^r\binom{n}{2r+1}x^{n-2r-1}(1-x^2)^{\frac{2r+1}{2}}, \quad \text{...(7)}$$

where $K=1.3.5\ldots(2n-1)$.

In (7) let $x=\cos\theta$ where $0<\theta<\frac{\pi}{2}$; then

$$\frac{d^{n-1}.(1-x^2)^{n-\frac{1}{2}}}{dx^{n-1}}=(-1)^{n-1}\frac{1.3.5\ldots(2n-1)}{n}\sin n\theta, \quad \text{.........(8)}$$

by using the expression for $\sin n\theta$

$$n\cos^{n-1}\theta\sin\theta-\binom{n}{3}\cos^{n-3}\theta\sin^3\theta+\binom{n}{5}\cos^{n-5}\theta\sin^5\theta-\ldots.$$

Formula (8) was first given by O. Rodrigues in 1815, but is usually attributed to Jacobi who, no doubt without knowledge of Rodrigues' work, published it in 1826. (See *Exercises* IV, 14, for another solution.)

Ex. 2. If $y=f(u)$ and $u=\log x$ find $\dfrac{d^ny}{dx^n}$.

The formula (1) is not suitable when $u=\log x$ and it is better to start afresh. A little consideration will show that the form to be tested is the following:

$$\frac{d^ny}{dx^n}=\frac{1}{x^n}\left\{f^{(n)}(u)-C_{n,1}f^{(n-1)}(u)+C_{n,2}f^{(n-2)}(u)-\ldots\right\}$$
$$=\frac{1}{x^n}\sum_{r=0}^{n-1}(-1)^rC_{n,r}f^{(n-r)}(u), \qquad C_{n,0}=1. \quad \text{....................(9)}$$

A further differentiation shows that the form is correct.

Now $C_{n,r}$ is independent of the form of $f(u)$, and to determine these coefficients we take $f(u)$ equal to e^{-tu} where t is any constant. Thus

$$y=e^{-tu}=e^{-t\log x}=x^{-t},$$

and therefore

$$\frac{d^ny}{dx^n}=(-1)^nt(t+1)(t+2)\ldots(t+n-1)x^{-t-n},$$

$$f^{(n-r)}(u)=\frac{d^{n-r}.e^{-tu}}{du^{n-r}}=(-1)^{n-r}t^{n-r}e^{-tu}=(-1)^{n-r}t^{n-r}x^{-t},$$

so that, by substituting these values in (9), we have the identity

$$t(t+1)(t+2)\ldots(t+n-1)=\sum_{r=0}^{n-1}C_{n,r}t^{n-r}.$$

By equating the coefficients of t^{n-r} the value of $C_{n,r}$ is found. When n is not a large integer the values of $C_{n,r}$ can be picked out without much trouble, but there does not seem to be any convenient explicit formula.

37. Other Methods for Higher Derivatives. The formula (1) of the last article is cumbrous, and other methods are available that often lead to interesting results apart from the particular formula for the derivative. The formula suggested by a few differentiations may have coefficients that can be more conveniently dealt with than by the method used in the last article.

I. *Comparison of Expansions.*

$y=f(x)=e^{tx^2}$. This is a particular case of § 36, Example 1. Express $f(x+h)$ as a series in two different ways.

(i) $$f(x+h)=\sum_{n=0}^{\infty}\frac{h^n}{n!}f^{(n)}(x).$$

Next we have $$f(x+h)=f(x)\,.\,e^{2xth}\,.\,e^{th^2},$$

and if each of these two exponentials be expanded in powers of h and their product formed the coefficient, $u_n(x)$ say, of $h^n/n!$ will, when multiplied by $f(x)$, be equal to $f^{(n)}(x)$ in (i). Thus we have the second expansion,

(ii) $$f(x+h)=f(x)\,.\left\{\sum_{n=0}^{\infty}\frac{(2xt)^n h^n}{n!}\right\}.\left\{\sum_{n=0}^{\infty}\frac{t^n h^{2n}}{n!}\right\}$$
$$=f(x)\sum_{n=0}u_n(x)\frac{h^n}{n!},$$

where

$$u_n(x)=(2x)^n t^n+\frac{n(n-1)}{1}(2x)^{n-2}t^{n-1}$$
$$+\frac{n(n-1)(n-2)(n-3)}{1\,.\,2}(2x)^{n-4}t^{n-2}+\dots \quad \dots(1)$$

The last term is independent of x when n is even and contains the first power of x when n is odd. Hence

$$\frac{d^n\,.\,e^{tx^2}}{dx^n}=e^{tx^2}\,.\,u_n(x). \qquad .(2)$$

If $t=\sqrt{(-1)}=i$ this formula gives the derivatives of $\cos(x^2)$ and $\sin(x^2)$ since $e^{ix^2}=\cos(x^2)+i\sin(x^2)$ and the real and imaginary parts on the two sides of (2) may be equated.

We take the same example to illustrate another method.

II. *Use of a differential equation.*

The form of the nth derivative of e^{tx^2} is easily seen to be the product of e^{tx^2} and a polynomial $u_n(x)$; the polynomial

is of the nth degree in x (or $2x$), the exponents decrease by 2 and the coefficient of $(2x)^n$ is t^n. Thus we have, when $y=e^{tx^2}$,

$$D_x^n y = e^{tx^2} u_n(x), \qquad \text{......(2')}$$

where

$$u_n(x) = c_0(2x)^n + c_1(2x)^{n-2} + \ldots + c_r(2x)^{n-2r} + \qquad \text{......}(\alpha)$$

and $c_0 = t^n$.

We now find a differential equation for $u_n(x)$. Write the value of Dy in the form

$$Dy = 2xty, \qquad \text{......(3)}$$

and differentiate this equation n times, using Leibniz's Theorem; then

$$D^{n+1}y = 2xt\, D^n y + 2nt\, D^{n-1}y$$

But

$$D^{n+1}y = e^{tx^2}u_{n+1},\ D^{n-1}y = e^{tx^2}u_{n-1};$$

therefore

$$u_{n+1} = 2xt\, u_n + 2nt\, u_{n-1}. \qquad \text{......(4)}$$

It would be possible to calculate u_{n+1} from (4) if u_n and u_{n-1} were known; now u_1 and u_2 are easily found so that $u_3, u_4, \ldots$ could be calculated. It is better, however, to find a differential equation for u_n; the process is a little troublesome though not really hard if the principle be grasped. We have in fact to eliminate u_{n-1} and u_{n+1} and put in their place u_n' and u_n'' where $u_n' = D_x u_n$, $u_n'' = D_x^2 u_n$.

Differentiate the equation (2′) once; therefore

$$D^{n+1}y = e^{tx^2} \cdot u_n' + 2xt\, e^{tx^2} \cdot u_n,$$

so that

$$u_{n+1} = u_n' + 2xt\, u_n. \qquad \text{......(5)}$$

Elimination of u_{n+1} between (5) and (4) gives

$$u_n' = 2nt\, u_{n-1}. \qquad \text{......(6)}$$

Differentiate (6) and for u_{n-1}' put $2(n-1)t\, u_{n-2}$, the value obtained from (6) by changing n into $n-1$; then

$$u_n'' = 2nt \cdot 2(n-1)t\, u_{n-2}. \qquad \text{......(7)}$$

Next in (4) put $n-1$ for n; therefore

$$u_n = 2x\, tu_{n-1} + 2(n-1)t\, u_{n-2}, \qquad \text{......(4')}$$

and then the elimination of u_{n-1} and u_{n-2} between (4′), (6) and (7) gives the required differential equation

$$u_n'' + 2xt\, u_n' - 2nt\, u_n = 0, \qquad \text{......(8)}$$

We now substitute in (8) the value of u_n given by (α); the equation must be then satisfied identically, and therefore the

coefficient of each power of x must be zero. The coefficient of $(2x)^{n-2r}$ is

$$[-2nt+2(n-2r)t]c_r+4(n-2r+2)(n-2r+1)c_{r-1},$$

and as this coefficient must be zero we find

$$c_r=\frac{(n-2r+2)(n-2r+1)}{rt}c_{r-1}, \quad r=1,\ 2,\ 3, \ldots$$

the coefficient of $(2x)^n$ being identically zero. Thus, since $c_0=t^n$ we get

$$c_r=\frac{n(n-1)(n-2)\ldots(n-2r+1)}{1\,.\,2\,.\ldots r}t^{n-r},$$

and therefore $u_n(x)$ is the same as was found by the first method.

When t is negative, say $t=-1$, the polynomial $u_n(x)$ has interesting properties.

Ex. 1. When $t=-1$ prove, by applying Rolle's Theorem, that the roots of $u_n(x)=0$ are all real and different.

Let $u_n(x)=v_n(x)$ when $t=-1$; then we take

$$f(x)=e^{-x^2}, \quad f^{(n)}(x)=e^{-x^2}v_n(x).$$

Now $f(x)=0$ for $x=-\infty$ and for $x=+\infty$; therefore $f'(x)$ vanishes for a value α of x. But $f'(\alpha)=e^{-\alpha^2}v_1(\alpha)$ and, as $e^{-\alpha^2}$ is not zero, $v_1(\alpha)=0$. Again, $f'(x)=0$ for the values $-\infty$, α, $+\infty$ of x, and therefore the derivative of $f'(x)$, that is, $f''(x)$ vanishes for a value, β say, between $-\infty$ and α and also for a value, γ say, between α and $+\infty$. As before $v_2(x)=0$ for $x=\beta$ and $x=\gamma$. Proceeding in this way it is readily seen that $v_n(x)$ is zero for n different values of x. It is besides clear from equation (8) that if $v_n(x)$ had two equal roots, each equal to λ say, we should have $v_n(\lambda)=0$, $v_n'(\lambda)=0$, and therefore also $v_n''(\lambda)=0$; if the equation be differentiated once it will be seen that we should have $v_n'''(\lambda)=0$, and so on, so that every derivative of $v_n(x)$ would vanish for $x=\lambda$ which is impossible since $v_n(x)$ is not identically zero.

Rodrigues' Formula. The following relation between derivatives, known as **Rodrigues' Formula**, is of importance in the theory of Legendre's Coefficients:

$$\frac{d^{n-r}\,.\,(x^2-1)^n}{dx^{n-r}}=\frac{(n-r)!}{(n+r)!}(x^2-1)^r\frac{d^{n+r}\,.\,(x^2-1)^n}{dx^{n+r}}.$$

The proof may be easily given by the method I. The function $\{2x+h+(x^2-1)h^{-1}\}^n$, where n is a positive integer, is not altered when $(x^2-1)h^{-1}$ is substituted in place of h. Now

$$\{2x+h+(x^2-1)h^{-1}\}^n=\frac{1}{h^n}\{((x+h)^2-1)\}^n,$$

and therefore by Taylor's Theorem is equal to

$$\frac{1}{h^n}\sum_{s=0}^{2n}\frac{h^s}{s!}\frac{d^s\,.\,(x^2-1)^n}{dx^s}=\sum_{s=0}^{2n}\frac{h^{s-n}}{s!}\frac{d^s\,.\,(x^2-1)^n}{dx^s}.$$

If for h we substitute $(x^2-1)h^{-1}$ the expansion becomes

$$\sum_{s=0}^{2n}\frac{(x^2-1)^{s-n}h^n}{s!}\cdot\frac{d^s\,.\,(x^2-1)^n}{dx^s}.$$

Since the two expansions last written are identical the coefficient of h^r is the same in both; for the second expansion $s=n-r$ and for the first $s=n+r$. Hence

$$\frac{(x^2-1)^{-r}}{(n-r)!}\frac{d^{n-r}\,.\,(x^2-1)^n}{dx^{n-r}}=\frac{1}{(n+r)!}\frac{d^{n+r}\,.\,(x^2-1)^n}{dx^{n+r}}$$

which at once gives the formula stated.

Ex. 2. If $P_n(x)=\dfrac{1}{2^n\,.\,n!}\dfrac{d^n\,.\,(x^2-1)^n}{dx^n}$,

prove that $P_n(x)$ satisfies the differential equation

$$\frac{d}{dx}\left\{(x^2-1)\frac{dy}{dx}\right\}=n(n+1)y,$$

or
$$(1-x^2)D^2y-2xDy+n(n+1)y=0.$$

Let $y=2^n\,.\,n!\,P_n(x)$; then Rodrigues' Formula $(r=1)$ gives

$$(x^2-1)\frac{dy}{dx}=(x^2-1)\frac{d^{n+1}\,.\,(x^2-1)^n}{dx^{n+1}}=n(n+1)\frac{d^{n-1}\,.\,(x^2-1)^n}{dx^{n-1}}$$

and therefore

$$\frac{d}{dx}\left\{(x^2-1)\frac{dy}{dx}\right\}=n(n+1)\frac{d^n\,.\,(x^2-1)^n}{dx^n}=n(n+1)y.$$

EXERCISES IV.

1. If u and v are functions of x prove that

$$vD^nu=D^n(uv)-\binom{n}{1}D^{n-1}(uDv)+\binom{n}{2}D^{n-2}(uD^2v)-\ldots$$
$$+(-1)^r\binom{n}{r}D^{n-r}(uD^rv)+\ldots+(-1)^nuD^nv.$$

2. If $y=(x^2+a^2)^{-1}$ and $x=a\cot\theta$ prove that

$$\frac{d^ny}{dx^n}=(-1)^n\frac{n!}{a^{n+2}}(\sin\theta)^{n+1}\sin(n+1)\theta,$$

and deduce that the nth derivative of $\tan^{-1}(x/a)$ is

$$(-1)^{n-1}(n-1)!\,a^{-n}\sin^n\theta\sin n\theta.$$

3. If $y=x(x^2+a^2)^{-1}$ and $x=a\cot\theta$ show that

$$\frac{d^ny}{dx^n}=(-1)^n\frac{n!}{a^{n+1}}(\sin\theta)^{n+1}\cos(n+1)\theta.$$

4. If $y=f(u)$ and $u=1/x$ prove that

$$(-1)^n\frac{d^ny}{dx^n}=\sum_{r=0}^{n-1}\frac{(n-1)!}{(n-r-1)!}\binom{n}{r}\frac{f^{(n-r)}(u)}{x^{2n-r}},$$

and show that

$$\frac{d^n\,.\,e^{a/x}}{dx^n}=(-1)^n\frac{e^{a/x}}{x^n}\sum_{r=0}^{n-1}\frac{(n-1)!}{(n-r-1)!}\binom{n}{r}\left(\frac{a}{x}\right)^{n-r}.$$

5. $$\frac{d^n\left(x^{n-1}e^{\frac{1}{x}}\right)}{dx^n}=(-1)^n\frac{e^{\frac{1}{x}}}{x^{n+1}}.$$ (Halphen.)

6. If $y=f(u)$ and $u=\sqrt{x}$ prove either by the general formula or independently that

$$\frac{d^ny}{dx^n}=\sum_{r=0}^{n-1}(-1)^r\frac{(n+r-1)!}{r!\,(n-r-1)!}\frac{f^{(n-r)}(u)}{(2\sqrt{x})^{n+r}},$$

and show that if $y=(1+a\sqrt{x})^{2n-1}$,

$$\frac{d^ny}{dx^n}=\frac{1\,.\,3\,.\,5\,\ldots\,(2n-1)}{2^n}\cdot\frac{a}{\sqrt{x}}\cdot\left(a^2-\frac{1}{x}\right)^{n-1}.$$

[Schlömilch gives the following proof, *Compendium* II, pp. 7, 8.

Let $\varrho=xt$ and $w=\sqrt{(1+t)}-1$; then § 36, (5) gives

$$A_{n,r}=\frac{1}{(\sqrt{x})^{2n-r}}\Big[D_t^n(w^r)\Big]_{t=0}.$$

Now $(t-w)\,dw/dt=\frac{1}{2}w$ and therefore, multiplying by w^{r-2},

$$\frac{1}{r-1}tD_t(w^{r-1})-\frac{1}{r}D_t(w^r)=\tfrac{1}{2}w^{r-1}.$$

Differentiate $(n-1)$ times as to t and then let $t=0$; thus we find the reduction-formula

$$\Big[D_t^n(w^r)\Big]_{t=0}=\frac{r(2n-r-1)}{2(r-1)}\Big[D_t^{n-1}(w^{r-1})\Big]_{t=0},$$

or, if $n-r$ is put in place of r so as to obtain $A_{n,\,n-r}$,

$$\Big[D_t^n(w^{n-r})\Big]_{t=0}=\frac{(n-r)(n+r-1)}{2(n-r-1)}\Big[D_t^{n-1}(w^{n-r-1})\Big]_{t=0}.$$

Now apply this formula till the index of w becomes unity and note that

$$\Big[D_t^{r+1}w\Big]_{t=0}=(-1)^r\frac{1\,.\,3\,.\,5\,\ldots\,(2r-1)}{2^{r+1}}.$$

The verification of the value of d^ny/dx^n is simple.]

7. If $y=f(u)$ and $u=e^x$ prove that

$$\frac{d^ny}{dx^n}=\sum_{r=1}^{n}\frac{c_re^{rx}}{r!}f^{(r)}(u),$$

where $$c_r=\sum_{s=0}^{r-1}(-1)^s\binom{r}{s}(r-s)^n.$$

8. If $y=(u+1)^{-1}$ and $u=e^x$ prove that

$$\frac{d^ny}{dx^n}=\frac{1}{e^x+1}\sum_{r=1}^{n}(-1)^r c_r\left(\frac{e^x}{e^x+1}\right)^r,$$

where c_r is given in Example 7. Show independently that

$$(-1)^n(e^x+1)^{n+1}\frac{d^ny}{dx^n}=a_ne^{nx}+a_{n-1}e^{(n-1)x}+\ldots+a_1e^x,$$

where
$$a_n=1^n,\ a_{n-1}=-\left(2^n-\frac{n+1}{1}1^n\right),$$

$$a_{n-2}=3^n-\frac{n+1}{1}\cdot 2^n+\frac{(n+1)n}{1\cdot 2}\cdot 1^n,\ \ldots .$$

9. If $y=(1-x^2)^{-\frac{1}{2}}$ show that

$$\frac{d^ny}{dx^n}=\frac{u_n(x)}{(1-x^2)^{n+\frac{1}{2}}},$$

where $u_n(x)$ is a polynomial in x of the nth degree in which the coefficient of x^n is $n!$ and the exponents of x decrease by 2. Establish the relations:

(i) $u_{n+1}-(2n+1)xu_n-n^2(1-x^2)u_{n-1}=0$;

(ii) $u_{n+1}=(1-x^2)u_n'+(2n+1)xu_n$;

(iii) $(1-x^2)u_n''+(2n-1)xu_n'-n^2u_n=0$,

and find $u_n(x)$. (Compare § 37, II.)

10. If $y=(1+x^2)^{-\frac{1}{2}}$ deduce from example 9 or prove independently

that
$$\frac{d^ny}{dx^n}=(-1)^n\frac{v_n(x)}{(1+x^2)^{n+\frac{1}{2}}},$$

where
$$(1+x^2)v_n''-(2n-1)xv_n'+n^2v_n=0.$$

Prove that the roots of $v_n(x)=0$ are all real and different.

11. Deduce from Example 10 the nth derivative of $\log\{x+\sqrt{(1+x^2)}\}$.

12. If $y=(x\log x)^n$ show that

$$\frac{1}{n!}\frac{d^ny}{dx^n}=1+S_1\log x+\frac{S_2}{2!}(\log x)^2+\ldots+\frac{S_n}{n!}(\log x)^n,$$

where S_r is the sum of the products, r at a time, of the numbers $1, 2, 3, \ldots, n$.

13. If $u=[f'(x)]^{-\frac{1}{2}}$ and $v=f(x)[f'(x)]^{-\frac{1}{2}}$ prove that

$$\frac{1}{u}\frac{d^2u}{dx^2}=\frac{1}{v}\frac{d^2v}{dx^2}.$$ (Goursat.)

14. If $z=(1-x^2)^{n-\frac{1}{2}}$ and $y=\dfrac{d^{n-1}.(1-x^2)^{n-\frac{1}{2}}}{dx^{n-1}}$ prove that

(i) $(1-x^2)\dfrac{dz}{dx}+(2n-1)xz=0$;

(ii) $(1-x^2)\dfrac{d^2y}{dx^2}-x\dfrac{dy}{dx}+n^2y=0$;

and that if $x=\cos\theta$ the equation (ii) becomes

$$\frac{d^2y}{d\theta^2}+n^2y=0.$$

Next show that when $x=1$

$$y=0 \text{ and } y(1-x^2)^{-\frac{1}{2}}=(-1)^{n-1}\,1\,.\,3\,.\,5 \qquad (2n-1),$$

and deduce that

$$y=(-1)^{n-1}\frac{1\,.\,3\,.\,5\,\ldots\,(2n-1)}{n}\sin n\theta.$$

15. If $z=(x^2-1)^n$ and $y=\frac{1}{2^n\,.\,n!}\frac{d^n\,.\,(x^2-1)^n}{dx^n}$ prove that

$$\text{(i)}\ \ (x^2-1)\frac{dz}{dx}-2nxz=0\,;$$

$$\text{(ii)}\ \ (1-x^2)\frac{d^2y}{dx^2}-2x\frac{dy}{dx}+n(n+1)y=0.$$

Show that $y=1$ when $x=1$ and $y=(-1)^n$ when $x=-1$ and deduce that the roots of the equation $y=0$ are all real and different and lie between -1 and $+1$. (See § 37, Ex. 1.)

16. If P and Q are two rational integral functions of x (polynomials) such that

$$\sqrt{(1-P^2)}=Q\sqrt{(1-x^2)}$$

prove that

$$\frac{dP}{dx}=n\sqrt{\frac{1-P^2}{1-x^2}},$$

where n is an integer.

[$$1-P^2=Q^2(1-x^2) \qquad \text{(i)}$$

so that, by differentiation,

$$-2PP'=2\{QQ'(1-x^2)-xQ\} \qquad \text{(ii)}.$$

From (i) Q is prime to P and therefore from (ii) Q is a factor of P'; then compare the coefficients of the highest powers in P^2 and Q^2, and also in P' and Q.]

38. Derivative of a Determinant. The proof of the rule for forming the derivative of a determinant of the nth order whose elements are functions of a variable x will be understood by consideration of a determinant D of the third order, say

$$D=\begin{vmatrix} & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ & & c_3 \end{vmatrix}$$

Let $\delta a_1\,\ldots,\,\delta c_3$ and δD be the increments of $a_1,\,\ldots,\,c_3$ and D corresponding to the increment δx of x; the determinant $D+\delta D$ is

$$\begin{vmatrix} a_1+\delta a_1 & a_2+\delta a_2 & a_3+\delta a_3 \\ b_1+\delta b_1 & b_2+\delta b_2 & b_3+\delta b_3 \\ c_1+\delta c_1 & c_2+\delta c_2 & c_3+\delta c_3 \end{vmatrix}$$

and may be expressed as the sum of 8 determinants, namely:

(i) the determinant D;

(ii) 3 determinants, each containing one column of increments,

$$\begin{vmatrix} \delta a_1 & a_2 & a_3 \\ \delta b_1 & b_2 & b_3 \\ \delta c_1 & c_2 & c_3 \end{vmatrix}, \quad \begin{vmatrix} a_1 & \delta a_2 & a_3 \\ b_1 & \delta b_2 & b_3 \\ c_1 & \delta c_2 & c_3 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_2 & \delta a_3 \\ b_1 & b_2 & \delta b_3 \\ c_1 & c_2 & \delta c_3 \end{vmatrix}.$$

(iii) 3 determinants, each containing two columns of increments, of the type

$$\begin{vmatrix} \delta a_1 & \delta a_2 & a_3 \\ \delta b_1 & \delta b_2 & b_3 \\ \delta c_1 & \delta c_2 & c_3 \end{vmatrix}.$$

(iv) 1 determinant, containing increments alone.

δD is the sum of the 7 determinants (ii), (iii) and (iv); when each of these determinants is divided by δx * and δx made to tend to zero, the determinants in (iii) will have each one column that tends to zero, and the determinant in (iv) two columns that tend to zero. Hence if accents indicate derivatives with respect to x we find

$$D' = \begin{vmatrix} a_1' & a_2 & a_3 \\ b_1' & b_2 & b_3 \\ c_1' & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & a_2' & a_3 \\ b_1 & b_2' & b_3 \\ c_1 & c_2' & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & a_2 & a_3' \\ b_1 & b_2 & b_3' \\ c_1 & c_2 & c_3' \end{vmatrix}.$$

If D were of the nth order the determinant (i) would be D while (ii) would contain n determinants each containing one column of increments,—the first column of the first determinant, the second column of the second determinant, ..., and the nth column of the nth determinant. All the remaining determinants, namely $(2^n - 1 - n)$ determinants, would contain at least two columns of increments and would therefore tend to zero when δx tends to zero. Hence the rule:

The derivative of a determinant of the nth order is the sum of n determinants which are obtained by substituting in turn in place of the elements in the 1st, 2nd, ..., nth columns the derivatives of the elements in the 1st, 2nd, ..., nth columns. Instead of "columns" the word "rows" may be used since a determinant is not altered by the interchange of rows and columns.

* To divide one of these determinants by δx, divide any one column of increments in it by δx.

Ex.
$$D = \begin{vmatrix} a-x, & h, & g \\ h, & b-x, & f \\ g, & f, & c-x \end{vmatrix}$$

$$D' = \begin{vmatrix} -1 & 0 & 0 \\ h, & b-x, & f \\ g, & f, & c-x \end{vmatrix} + \begin{vmatrix} a-x, & h, & g \\ 0 & -1 & 0 \\ g, & f, & c-x \end{vmatrix} + \begin{vmatrix} a-x, & h, & g \\ h, & b-x, & f \\ 0 & 0 & -1 \end{vmatrix}$$
$$= -3x^2 + 2(a+b+c)x - (bc+ca+ab-f^2-g^2-h^2).$$

39. Linear Dependence of Functions. Let $f_1, f_2, \dots, f_n$ be n functions of x defined for a range $a \leqq x \leqq b$; if there are n constants $c_1, c_2, \dots, c_n$, not all zero, such that

$$c_1 f_1 + c_2 f_2 + \dots + c_n f_n = 0, \qquad \text{.....................(1)}$$

the functions $f_1, f_2, \dots, f_n$ are said to be *linearly dependent*; if there is no such set of constants the functions are said to be *linearly independent*.

Obviously if $c_1, c_2, \dots, c_n$ is one set of constants that satisfies (1) so is the set $kc_1, kc_2, \dots, kc_n$ where k is any constant that is not zero. Hence one of the constants may be taken to be ± 1.

The test now to be given for the linear dependence requires that each of the functions should have all the derivatives up to and including the $(n-1)$th.

First. Suppose that $f_1, f_2, \dots, f_n$ are linearly dependent and that equation (1) therefore holds. The equation (1) may be differentiated $(n-1)$ times; differentiate and let accents denote derivatives with respect to x. We thus find the following n equations, including equation (1):

$$\begin{array}{lllll} c_1 f_1 & + c_2 f_2 & + \dots + c_{n-1} f_{n-1} & + c_n f_n & = 0, \\ c_1 f_1' & + c_2 f_2' & + \dots + c_{n-1} f_{n-1}' & + c_n f_n' & = 0, \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \\ c_1 f_1^{(n-2)} & + c_2 f_2^{(n-2)} & + \dots + c_{n-1} f_{n-1}^{(n-2)} & + c_n f_n^{(n-2)} & = 0, \\ c_1 f_1^{(n-1)} & + c_2 f_2^{(n-1)} & + \dots + c_{n-1} f_{n-1}^{(n-1)} & + c_n f_n^{(n-1)} & = 0. \end{array} \qquad \text{..........(2)}$$

Now the coefficients $c_1, c_2, \dots, c_n$ are not all zero and therefore the determinant W of the equations must be zero for every x such that $a \leqq x \leqq b$ where

$$W = \begin{vmatrix} f_1 & f_2 & \dots f_{n-1} & f_n \\ f_1' & f_2' & \dots f_{n-1}' & f_n' \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ f_1^{(n-2)} & f_2^{(n-2)} & \dots f_{n-1}^{(n-2)} & f_n^{(n-2)} \\ f_1^{(n-1)} & f_2^{(n-1)} & \dots f_{n-1}^{(n-1)} & f_n^{(n-1)} \end{vmatrix} \qquad \text{.....................(3)}$$

W is called the *Wronskian* of the functions $f_1, f_2, \dots, f_n$, and is denoted more fully as $W(f_1, f_2, \dots, f_n)$.

Hence the condition $W=0$ is a *necessary* condition for the linear dependence of $f_1, f_2, \dots, f_n$ and it is to be noted that W must be *identically zero*, that is, zero for every value of x in the interval (a, b).

Second. The condition $W=0$, which has been seen to be necessary, is also a *sufficient* condition for the linear dependence of $f_1, f_2, \dots, f_n$ *provided* the Wronskian, W_1 say, of the $(n-1)$ functions $f_1, f_2, \dots, f_{n-1}$ is not zero for $a \leqq x \leqq b$.

Consider the system of $(n-1)$ equations:

$$\begin{aligned}
c_1 f_1 &+ c_2 f_2 + \dots + c_{n-1} f_{n-1} = f_n, \\
c_1 f_1' &+ c_2 f_2' + \dots + c_{n-1} f_{n-1}' = f_n', \\
c_1 f_1'' &+ c_2 f_2'' + \dots + c_{n-1} f_{n-1}'' = f_n'', \qquad \dots\dots\dots\dots\dots(4) \\
&\dots\dots\dots\dots \\
c_1 f_1^{(n-2)} &+ c_2 f_2^{(n-2)} + \dots + c_{n-1} f_{n-1}^{(n-2)} = f_n^{(n-2)}.
\end{aligned}$$

The determinant of this system is the Wronskian W_1 and is therefore not zero. Hence the system determines $c_1, c_2, \dots, c_{n-1}$ and these numbers will usually be functions of x; it has to be shown that if W vanishes for every x in the interval (a, b) the numbers $c_1, c_2, \dots, c_{n-1}$ will be constants and then the first of equations (4) proves the linear dependence of $f_1, f_2, \dots, f_n$.

Let it be first noted that if $W=0$ we can add the following equation to (4), namely:

$$c_1 f_1^{(n-1)} + c_2 f_2^{(n-1)} + \dots + c_{n-1} f_{n-1}^{(n-1)} = f_n^{(n-1)}. \qquad \dots\dots\dots\dots(4')$$

For in virtue of (4) we have

$$W = \begin{vmatrix}
f_1, & f_2, & \dots f_{n-1}, & 0 \\
f_1', & f_2', & \dots f_{n-1}', & 0 \\
\dots & \dots & \dots & \dots \\
f_1^{(n-2)}, & f_2^{(n-2)} & \dots f_{n-1}^{(n-2)}, & 0 \\
f_1^{(n-1)}, & f_2^{(n-1)} & \dots f_{n-1}^{(n-1)}, & f_n^{(n-1)} - c_1 f_1^{(n-1)} - c_2 f_2^{(n-1)} \dots - c_{n-1} f_{n-1}^{(n-1)}
\end{vmatrix}$$

$$= W_1 \times (f_n^{(n-1)} - c_1 f_1^{(n-1)} - c_2 f_2^{(n-1)} - \dots - c_{n-1} f_{n-1}^{(n-1)})$$

and $W=0$, $W_1 \neq 0$ so that equation (4′) follows.

Now differentiate the first equation in (4); therefore

$$\begin{aligned}
&c_1' f_1 + c_2' f_2 + \dots + c_{n-1}' f_{n-1} \\
&\quad + c_1 f_1' + c_2 f_2' + \dots + c_{n-1} f_{n-1}' = f_n',
\end{aligned}$$

which by the second equation in (4) reduces to

$$c_1'f_1 + c_2'f_2 + \dots + c_{n-1}'f_{n-1} = 0.$$

If the other equations in (4) be differentiated it will be seen in the same way that each equation reduces to an equation in $c_1', c_2', \dots, c_{n-1}'$, the right-hand side being zero; for the last equation in (4) this result follows from equation (4′). Hence we have the set of $(n-1)$ homogeneous equations

$$\begin{aligned} c_1 f_1 \quad &+ c_2'f_2 \quad + \dots + c_{n-1}'f_{n-1} = 0, \\ c_1'f' \quad &+ c_2'f_2' \quad + \dots + c_{n-1}'f_{n-1}' = 0, \\ & \\ c_1'f_1^{(n-2)} &+ c_2'f_2^{(n-2)} + \dots + c_{n-1}'f_{n-1}^{(n-2)} = 0. \end{aligned}$$

The determinant of the system is W_1 which is not zero and therefore each of the numbers $c_1', c_2', \dots, c_{n-1}'$ is zero, so that each of the numbers $c_1, c_2, \dots, c_{n-1}$ is constant. Hence, by (4),

$$f_n = c_1 f_1 + c_2 f_2 + \dots + c_{n-1} f_{n-1},$$

and the functions $f_1, f_2, \dots, f_n$ are linearly dependent.

Cor. If W and W_1 are identically zero but W_2, the Wronskian of the $(n-2)$ functions $f_1, f_2, \dots, f_{n-2}$, not zero, then, by what has been proved, the $(n-1)$ functions $f_1, f_2, \dots, f_{n-1}$ are linearly dependent and there are therefore $(n-1)$ constants $c_1, c_2, \dots, c_{n-1}$, not all zero, such that

$$c_1 f_1 + c_2 f_2 + \dots + c_{n-1} f_{n-1} = 0.$$

The n functions $f_1, f_2, \dots, f_n$ are therefore linearly dependent because in (1) we may make c_n zero and the n constants $c_1, c_2, \dots, c_n$ are not *all* zero.

Similarly, there is linear dependence of the n functions if W, W_1 and W_2 are identically zero but W_3, the Wronskian of the $(n-3)$ functions $f_1, f_2, \dots, f_{n-3}$ not zero, and so on.

Ex. 1. The functions e^x, xe^x, x^2e^x aro linearly independent.

Here,

$$W = \begin{array}{lll} & xe^x, & x^2e^x \\ & (x+1)e^x, & (x^2+2x)e^x \\ {}^x, & (x+2)e^x, & (x^2+4x+2)e^x \end{array} = 2e^{3x}$$

Ex. 2. The functions $\sin x$, $\cos x$, $\sin(x+\alpha)$ are linearly dependent.

Here, $\quad W = 0. \quad c_1 \sin x + c_2 \cos x + c_3 \sin(x+\alpha) = 0$

if $\quad c_1 = -\cos\alpha, \; c_2 = -\sin\alpha, \; c_3 = 1.$

Ex. 3. Show that the derivative of a Wronskian is obtained by differentiating each element of the last row.

Ex. 4. If $y_1, y_2, \dots, y_n$ are functions of x and if x is a function of t prove that

$$W_x(y_1, y_2, \dots, y_n) = \left(\frac{dt}{dx}\right)^{\frac{1}{2}n(n-1)} W_t(y_1, y_2, \dots, y_n),$$

where the suffixes x and t denote that the derivatives in the Wronskians are derivatives with respect to x and t respectively.

Ex. 5. If $y, y_1, y_2, \dots, y_n$ are all functions of x show that

$$W_x(yy_1, yy_2, yy_3, \dots, yy_n) = y^n W_x(y_1, y_2, y_3, \dots, y_n).$$

Ex. 6. If y is a function of x prove that

$$W(1, 2y, 3y^2, \dots, ny^{n-1}) = n!\,(n-1)!\,(n-2)! \dots 2!\,1 \,.\left(\frac{dy}{dx}\right)^{\frac{1}{2}n(n-1)}.$$

CHAPTER IV

FUNCTIONS OF SEVERAL VARIABLES. DERIVATIVES. DIFFERENTIALS. CHANGE OF VARIABLES

40. Functions of more than one Variable. The characteristic properties of a function of n independent variables may usually be understood by the study of a function of two or of three variables and unless some definite purpose is to be served the restriction to not more than three independent variables will be generally maintained; this restriction has the considerable advantage of simplifying the formulae and reducing the mere mechanical labour.

By extension of the usage of analytical geometry a set of values $a_1, a_2, \dots, a_n$ of n variables will often be called "the point $(a_1, a_2, \dots, a_n)$." The set of values $x_1, x_2, \dots, x_n$ other than $a_1, a_2, \dots, a_n$ that satisfy the conditions

$$|x_1 - a_1| < \varrho, |x_2 - a_2| < \varrho, \dots, |x_n - a_n| < \varrho,$$

where ϱ is an arbitrarily small positive number, is said to form a "neighbourhood" of the point $(a_1, a_2, \dots, a_n)$. The neighbourhood may, however, be specified in other, though equivalent, ways; for example, the points inside the sphere $x^2 + y^2 + z^2 = \varrho^2$ may be taken as the point $(0, 0, 0)$ and its neighbourhood.

The set of values of the variables for which a function is defined is called "the region (or, domain) of definition of the function." A function may be defined for integral values alone of its variables or for variables that vary continuously within given limits or that take all real values. The simplest type is the polynomial which is defined for all real values; next in simplicity is the quotient of two polynomials which is defined for all real values except such as make the divisor

zero. In general, all the usual functions of a single variable reappear.

The language and conceptions of geometry necessarily play an important part and certain assumptions are made that may be illustrated by considering a function $f(x, y)$ of two independent variables. The function may be defined for the whole plane or for a part or parts of the plane that are bounded by closed curves. It is supposed that a closed curve C, *without double points*, divides the plane into two regions, an interior and an exterior, such that any two points in any one region can be joined by a path that lies wholly within that region while every path that joins a point of one region to a point of the other cuts the curve C that separates the two regions.

If $x = \varphi(t)$, $y = \psi(t)$, where φ and ψ are continuous functions of t, are the equations of a curve the curve is *closed* when φ and ψ are periodic functions of t with period ω. In this case the points "t" and "$t + n\omega$," where n is any integer, are identical. If on the other hand t' and t'' are two values of t *which do not differ by ω or a multiple of ω*, such that $\varphi(t') = \varphi(t'')$ and $\psi(t') = \psi(t'')$ the point t' or t'' is *a double point*.

A curve of the kind spoken of may be a circle or an ellipse or any "ordinary" curve that does not intersect itself (like a lemniscate), but it may equally well be a rectangle or polygon, and the path spoken of as joining two points may be a "curve" in the ordinary sense or a "broken line" consisting, for example, of a set of segments that are alternately parallel to the coordinate axes.

The functions $\varphi(t)$ and $\psi(t)$ are supposed to be continuous but nothing is prescribed as to their derivatives; it may be said at once, however, that it will be assumed that every curve may be divided into a finite number of parts such that for each part $\varphi'(t)$ and $\psi'(t)$ exist.

When a function $f(x, y)$ is defined for the region bounded by a closed curve C the *region of definition* is said to be *closed* if $f(x, y)$ is defined for all points within and on the curve C, but *open* or *unclosed* when the function is defined for points within but not on the curve C.

The extension to regions for functions of three variables is fairly simple, and while geometry fails for the ordinary mortal when he enters regions of n dimensions it is possible at least to understand what is meant when it is said that the region is that within the sphere $x_1^2 + x_2^2 + \dots + x_n^2 = c^2$.

Limiting Points. If S is an infinite set of points lying in a region A the point $(a_1, a_2, \dots, a_n)$ is called a *limiting point* or *a point of condensation* of the set when an infinite number of points of the set lie in every neighbourhood of $(a_1, a_2, \dots, a_n)$;

the limiting point itself may or may not be a point of the set (§ 17).

If, for example, S consists of all the points inside a sphere, every point inside or on the surface of the sphere is a limiting point. The region A bounded by the sphere is "open" if it does not include the points on the surface—that is, if it does not include *all* its limiting points—but "closed" if it contains *all* its limiting points, because it then contains all points inside and on the surface of the sphere.

41. Limits and Continuity. A function $f(x_1, x_2, \ldots, x_n)$ of n independent variables $x_1, x_2, \ldots, x_n$ is said to tend to a limit l when $x_1, x_2, \ldots, x_n$ tend respectively to $a_1, a_2, \ldots, a_n$ if, given the arbitrarily small positive number ε, there is a positive number η such that

$$|f(x_1, x_2, \ldots, x_n) - l| < \varepsilon$$

when $|x_1 - a_1|, |x_2 - a_2|, \ldots, |x_n - a_n|$ are each less than η, the set of values $x_1 = a_1, x_2 = a_2, \ldots, x_n = a_n$ being excluded.

The modifications required when one or more of the variables tend to infinity or when l is $+\infty$ or $-\infty$ may be left for the student to state; with his previous work there should be no difficulty.

It should be specially noted, however, what the above definition implies; there must be no assumption of any relation between the variables as they tend to their respective limits. For example take $f(x, y)$ where

$$f(x, y) = 2xy/(x^2 + y^2).$$

If $x \to 0$, y being constant, $f(x, y) \to 0$ and if $y \to 0$, x being constant, $f(x, y) \to 0$ so that these limits of $f(x, y)$ exist and are the same when $x \to 0$ and when $y \to 0$. On the other hand $f(x, y)$ has no limit when x and y tend independently to zero; for if we put $x = r\cos\theta$, $y = r\sin\theta$, we see that $f(x, y) = \sin 2\theta$, so that near $(0, 0)$ $f(x, y)$ may take any value between -1 and 1 and therefore has no limit. It has to be noticed that 0 is not a value that x can take when $y = 0$ or that y can take when $x = 0$. The assumption that y is constant when $x \to 0$ is a violation of the conditions imposed by the definition, just as the assumption $y = x$ or any other relation between x and y would be.

Continuity. The function $f(x, y, z)$, the case of three variables being taken, is said to be continuous at the point (a, b, c) of a region for which it is defined if, given the arbitrarily small positive number ε, there is a positive number η such that $|f(x, y, z)-f(a, b, c)|<\varepsilon$ when $|x-a|$, $|y-b|$ and $|z-c|$ are each less than η.

It is necessary, therefore, for the continuity of $f(x, y, z)$ at (a, b, c) that $f(x, y, z)$ should tend to a limit when x, y, z tend to a, b, c respectively, and also that that limit should be the *value* $f(a, b, c)$ which, by hypothesis, exists since (a, b, c) is in the region for which $f(x, y, z)$ is defined. If (a, b, c) is on the boundary of the region the values of x, y, z that satisfy the conditions $|x-a|<\eta$, $|y-b|<\eta$ and $|z-c|<\eta$ must all, like $f(a, b, c)$ itself, be in the region of definition of the function.

A point to be particularly noticed is that $f(x, y, z)$ may be a continuous function of each variable when the other two are constant and yet not a continuous function of x, y and z. This peculiarity (for it does at first sight seem peculiar) is illustrated by the function $2xy/(x^2+y^2)$ just considered above. If $f(x, y)$ is defined to have the value zero at $(0, 0)$ it is defined for every neighbourhood of $(0, 0)$; as has been seen, $f(x, y)$ tends to its value 0 when $x\to 0$ and y is constant, or when $y\to 0$ and x is constant, and is therefore continuous at $(0, 0)$ when considered as a function of a single variable x or as one of a single variable y. On the other hand, $f(x, y)$ tends to no limit when the independent variables x and y tend to zero, and is therefore not continuous at $(0, 0)$.

42. Sequence of Decreasing Regions. The conceptions of infinite sets and of upper and lower bounds have no special restriction to functions of one variable, but the method of the decreasing interval used in the proof of various theorems requires a little explanation when it is applied to regions of two or more dimensions.

For definiteness, consider an *area* A bounded by a curve C; the principle is obviously applicable to regions of three or higher dimensions and the description is greatly facilitated by restriction of the region to a plane area.

The area will lie completely inside a rectangle R given by the equations $x=a$, $x=b>a$ and $y=c$, $y=d>c$. Let the rectangle R be divided

into four equal rectangles by the lines $x=\frac{1}{2}(a+b)$, $y=\frac{1}{2}(c+d)$, and let one of these four rectangles be selected, it being understood in this and all subsequent choice of rectangles that the selected rectangle contains a continuous piece of the area A. If the sides of this rectangle, R_1 say, are given by $x=a_1$, $x=b_1>a_1$ and $y=c_1$, $y=d_1>c_1$ then $a\leqq a_1<b_1\leqq b$, $c\leqq c_1<d_1\leqq d$ while $b_1-a_1=\frac{1}{2}(b-a)$, $d_1-c_1=\frac{1}{2}(d-c)$.

Next divide the rectangle R_1 into four equal rectangles by the lines $x=\frac{1}{2}(a_1+b_1)$, $y=\frac{1}{2}(c_1+d_1)$; select one of these four and let it be called R_2. The sides of R_2 will be given by $x=a_2$, $x=b_2>a_2$ and $y=c_2$, $y=d_2>c_2$, and the following relations will hold :

$$a\leqq a_1\leqq a_2<b_2\leqq b_1\leqq b,\quad b_2-a_2=\frac{1}{2^2}(b-a),$$

$$c\leqq c_1\leqq c_2<d_2\leqq d_1\leqq d,\quad d_2-c_2=\frac{1}{2^2}(d-c).$$

Proceeding in this way we obtain a sequence of decreasing rectangles (R_n). The sequences (a_n) and (c_n) are increasing and the sequences (b_n) and (d_n) decreasing sequences; $a_n<b_n$ and $c_n<d_n$ while

$$b_n-a_n=(b-a)/2^n,\ d_n-c_n=(d-c)/2^n.$$

Hence the sequences (a_n) and (b_n) determine a number ξ, the sequences (c_n) and (d_n) a number η and the point (ξ,η) is common to each rectangle R_n, each rectangle being closed (§ 16).

If the region were three-dimensional it might first be included in a cuboid (or rectangular parallelepiped) K bounded by the planes

$$x=a,\ x=a'\ ;\ y=b,\ y=b'\ ;\ z=c,\ z=c'\ ;\ a'>a,\ b'>b,\ c'>c.$$

The first step is to divide K into 8 cuboids by the planes

$$x=\tfrac{1}{2}(a+a'),\quad y=\tfrac{1}{2}(b+b'),\quad z=\tfrac{1}{2}(c+c'),$$

and to select one of these (call it K_1); its boundaries would be the planes

$$x=a_1,\ x=a_1'>a_1\ ;\ y=b_1,\ y=b_1'>b_1,\ z=c_1,\ z=c_1'>c_1$$

where $\quad a\leqq a_1<a_1'\leqq a',\quad b\leqq b_1<b_1'\leqq b',\quad c\leqq c_1<c_1'\leqq c'$

and $\quad a_1'-a_1=\frac{1}{2}(a'-a),\quad b_1'-b_1=\frac{1}{2}(b'-b),\quad c_1'-c_1=\frac{1}{2}(c'-c).$

Operate on K_1 in the same way and so on. The process determines a point (ξ,η,ζ) that is common to each cuboid K_n.

Here, and above, if the region is closed, the point found is a point of the region.

Notation. If P is the point (x,y,z) it is often convenient to denote the value of $f(x,y,z)$ at P by the symbol $f(P)$. A point $P'(x',y',z')$ is in a neighbourhood of P if

$$0<|x'-x|<\varrho,\ 0<|y'-y|<\varrho,\ 0<|z'-z|<\varrho,$$

or, if the length PP' is not zero but is less than ϱ. (One or two but not all of the differences $|x'-x|$, $|y'-y|$ and $|z'-z|$ may be zero.)

Theorems I and II of § 13 on the upper and lower bounds, M and m say, of a bounded set need no new proof, while Theorems I and II of § 15 on the limits of bounded monotonic functions are also valid for functions of several variables. Thus, if $f(x, y, z)$ increases (or does not decrease) when each of the variables x, y, z increases but is always less than a fixed number k, then $f(x, y, z)$ tends to a limit which is not greater than k when x, y, z tend to infinity.

Similarly Theorems I and II of § 17 need no new investigation. The important theorem of § 27 for a function $f(x, y, z)$ say, *namely that* " if M is the upper bound of $f(x, y, z)$ in a region R there is at least one point $P(\xi, \eta, \zeta)$ such that the upper bound is also M in any neighbourhood of P," may be proved at once by using a sequence of decreasing regions (R_n) instead of a sequence of decreasing intervals. The method is the same whether the sequence be a sequence of intervals or a sequence of regions.

The condition that $f(x, y, z)$ or $f(P)$ should tend to a limit when x, y, z tend respectively to ξ, η, ζ is that there should be a neighbourhood of the point A (ξ, η, ζ) such that $|f(P')-f(P'')|$ will be less than ε (where ε has the usual meaning) when P' and P'' are any two *admissible* points (x', y', z') and (x'', y'', z'') of that neighbourhood. [The point (x', y', z') is *admissible* when $f(x, y, z)$ is defined for the values x', y', z'.] In other language, there must be a positive number ϱ such that $|f(x', y', z')-f(x'', y'', z'')|<\varepsilon$ when each of the differences $|\xi-x'|$, $|\xi-x''|$, $|\eta-y'|$, $|\eta-y''|$, $|\zeta-z'|$, $|\zeta-z''|$ lies between 0 and ϱ. (Some but not all of these differences may be zero.)

That the condition is necessary is obvious. To prove the sufficiency of the condition proceed as in § 21. When it is satisfied we have

$$f(P'')-\varepsilon<f(P')<f(P'')+\varepsilon,$$

when P' and P'' are any two admissible points of the neighbourhood of A. The set of values $f(P')$ is therefore bounded and has maximum and minimum limits G and g, and so on, the rest of the proof being, except for verbal changes, the same as in § 21. (The meaning of the terms " upper limit of

indetermination" and "lower limit of indetermination" needs no further explanation.)

43. Theorems on Continuous Functions. The theorems of § 28 for functions of one variable are easily extended to functions of several variables; the theorems will now be stated, but the proofs will usually be indicated very briefly, if at all, since they are little more than repetitions of those for a function of one variable.

THEOREM I. *If $f(x, y, z)$ is continuous at (a, b, c) and if $f(a, b, c)$ is not zero, then $f(x, y, z)$ has the same sign as $f(a, b, c)$ at all points (x, y, z) in some neighbourhood of (a, b, c).*

THEOREM II. *If $f(x, y, z)$ is continuous at all points of a closed region R and if (x', y', z') and (x'', y'', z'') are two different points of the region at which $f(x, y, z)$ has two different values A and B then $f(x, y, z)$ takes in the region R all values between A and B.*

The method of § 28 may be adopted to prove the theorem, but the following method reduces the proof to that for a function of one variable. Assume, as will be proved in § 44 from the definition of continuity, that if $x=f_1(t)$, $y=f_2(t)$, $z=f_3(t)$, where f_1, f_2, f_3 are continuous functions of t, the function $f(x, y, z)$ becomes $F(t)$ where $F(t)$ is a continuous function of t. If t' and t'' give the points (x', y', z') and (x'', y'', z'') respectively, then $F(t)$ takes all values between A and B as t varies from t' to t''; if A and B have opposite signs there is at least one point τ or (ξ, η, ζ) at which the function $f(x, y, z)$ is zero. But, in general, there is an unlimited number of functions f_1, f_2, f_3—or, in geometrical language, an unlimited number of paths from (x', y', z') to (x'', y'', z'') that lie in the region R—and therefore $f(x, y, z)$ takes every value between A and B infinitely often.

The following theorem—the theorem of *uniform continuity*—will be proved for a function of two independent variables x and y but the method of proof is quite general. The method is not quite the same as that used in § 28 though not essentially different and the student might, as an exercise, apply the following method to the theorem of § 28.

THEOREM III. *If $f(x, y)$ is continuous at all points of a closed region A there is a positive number h such that, ε having the usual meaning, $| f(x', y') - f(x'', y'') | < \varepsilon$ where (x', y') and*

(x'', y'') are any two points in the region A, such that $|x'-x''|<h$ and $|y'-y''|<h$.

The proof is in two parts; the notation $f(P)$ to denote the value of $f(x, y)$ will be freely used.

(1) If ε_1 is any given arbitrarily small positive number the region A can be divided into a *finite* number of smaller regions (*sub-regions* they may be called) such that if P' and P'' are any two points in any sub-region $|f(P')-f(P'')|$ will be less than ε_1.

If such a division of the region A is impossible there will be, whatever division be made, at least one sub-region in which two points P' and P'' can be found such that $|f(P')-f(P'')| \geqq \varepsilon_1$; let such a sub-region be called, for convenience, *special*. It has now to be proved that A contains no special sub-region.

Let the area A be enclosed in a rectangle R, and, as in § 42, proceed to form a sequence (R_n) of rectangles. Of the four rectangles constructed at the first step one at least must contain a special sub-region, because if none of them did neither would A; select the rectangle (or one of the rectangles) that contains a special sub-region and call it R_1. Operate on R_1 in the same way and select a rectangle R_2 that contains a special sub-region, and so on. A sequence $R_1, R_2, \ldots$ of rectangles is thus obtained; the sequence determines a point P_0 of A such that *every region* within which P_0 lies contains a special sub-region. This conclusion will now be shown to be inconsistent with the continuity of $f(x, y)$.

Since $f(x, y)$ is continuous at (ξ, η) there is a region, σ say, within which $P_0(\xi, \eta)$ lies, such that if P' and P'' are *any two points* in σ

$$|f(P_0)-f(P')|<\tfrac{1}{2}\varepsilon_1, \quad |f(P_0)-f(P'')|<\tfrac{1}{2}\varepsilon_1,$$

and therefore $$|f(P')-f(P'')|<\varepsilon_1.$$

Now n may be taken so large (but finite) that the rectangle R_n will contain P_0 (either within or on its boundary) and lie wholly inside the region σ; therefore this rectangle R_n (a region that includes P_0) contains no special sub-region, since P' and P'' may be any two points in R_n.

Hence the hypothesis that the region A cannot be divided into a *finite* number of sub-regions of the kind stated is inadmissible.

(2) Theorem III now follows at once. Take $\varepsilon_1 = \frac{1}{2}\varepsilon$ and let the area A be covered by two sets of straight lines parallel to the axes of x and y, the distance h between two consecutive parallels being the same for each set. By (1) it is possible to choose h so that if P' and P'' are any two points in, or on the boundary of, a sub-region or square of side h we shall have

$$|f(P') - f(P'')| < \tfrac{1}{2}\varepsilon.$$

Now any two points $P'(x', y')$ and $P''(x'', y'')$ for which $|x' - x''| < h$ and $|y' - y''| < h$ must *either* lie in one and the same square in which case $|f(P') - f(P'')| < \frac{1}{2}\varepsilon < \varepsilon$ *or else* lie in adjacent squares. In this case if P is a point on the side common to the two squares (or if P is a common vertex of the two squares)

$$\begin{aligned}|f(P') - f(P'')| &= |f(P') - f(P) + f(P) - f(P'')| \\ &\leqq |f(P') - f(P)| + |f(P) - f(P'')| < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \text{ or } \varepsilon.\end{aligned}$$

Of course at the boundary of the region A the sub-regions will as a rule not be complete squares, but this makes no difference in the proof. Theorem III is therefore proved.

It thus follows that every function which is a continuous function of its variables, when these assume any values in a closed region, is a uniformly continuous function of its variables.

THEOREM IV. *If $f(x, y, z)$ is continuous at all points of a closed region it is bounded in that region.*

THEOREM V. *If $f(x, y, z)$ is continuous, and therefore bounded, at all points of a closed region there is at least one point (ξ, η, ζ) of the region for which $f(\xi, \eta, \zeta) = M$, the upper bound of $f(x, y, z)$, and at least one point (ξ', η', ζ') for which $f(\xi', \eta', \zeta') = m$, the lower bound of $f(x, y, z)$.*

The proofs of these two theorems may be left to the student as little more than verbal changes are needed to adapt the proofs for a function of a single variable.

44. Function of Functions. If $F(x_1, x_2, \dots, x_n)$ is a continuous function of the n variables $x_1, x_2, \dots, x_n$ in a region D and if the variables $x_1, x_2, \dots, x_n$ are continuous functions of m variables $y_1, y_2, \dots, y_m$ in a region D' the function $F(x_1, x_2, \dots, x_n)$, when expressed as a function $\varphi(y_1, y_2, \dots, y_m)$, is a continuous function of $y_1, y_2, \dots, y_m$ in the region D'.

Suppose $n=3$, $m=2$ and let the two sets of variables be denoted by x, y, z and s, t where

$$x=f(s, t),\quad y=g(s, t),\quad z=h(s, t)\ ;\ F(x, y, z)=\varphi(s, t)\ ;$$

also let $\qquad x'=f(s', t'),\ y'=g(s', t'),\ z'=h(s', t').$

Given ε as usual it is possible, since $f(x, y, z)$ is continuous, to choose $\eta(>0)$ so that

$$|\,f(x', y', z')-f(x, y, z)\,|<\varepsilon,$$

when $|\,x'-x\,|$, $|\,y'-y\,|$ and $|\,z'-z\,|$ are each less than η. Again since x, y, z are continuous functions of s and t it is possible to choose $\zeta(>0)$ so that

$$|\,f(s', t')-f(s, t)\,|<\eta,\quad |\,g(s', t')-g(s, t)\,|<\eta,$$
$$|\,h(s', t')-h(s, t)\,|<\eta$$

when $|\,s'-s\,|$ and $|\,t'-t\,|$ are each less than ζ. Hence $|\,\varphi(s', t')-\varphi(s, t)\,|<\varepsilon$ when $|\,s'-s\,|$ and $|\,t'-t\,|$ are each less than ζ and therefore $\varphi(s, t)$ is continuous in the region D'.

An important case of this theorem is that in which x, y, z are functions of a single variable t, so that the equations $x=f(t)$, $y=g(t)$, $z=h(t)$ can be taken as giving a curve; the function $F(x, y, z)$ or $\varphi(t)$ is therefore continuous for all points on the curve so long as the curve is within the region of definition. When x, y, z are functions of two variables the point (x, y, z) is restricted to a surface.

45. Partial Derivatives. Mean Value Theorem. In Chapter XI of the *Elementary Treatise* partial derivatives have been defined and various theorems proved; it seems desirable, however, to re-state briefly the fundamental equations and to present a more systematic treatment of the theory of differentials. The student is recommended to revise carefully §§ 92, 93 of that chapter which deal with the rate of variation in a given direction (important for its applications in mathematical physics) and with the interchangeability of the order of differentiation, the proof of which is often found difficult.

Suppose $f(x, y, z)$ and its partial derivatives f_x, f_y, f_z to be continuous; the increment δf corresponding to increments h, k, l in x, y, z respectively is given by the equation

$$\delta f=f(x+h, y+k, z+l)-f(x, y, z),$$

and this may be stated in the form

$$\delta f = [f(x+h,\ y+k,\ z+l) - f(x,\ y+k,\ z+l)]$$
$$+ [f(x, y+k, z+l) - f(x, y, z+l)] + [f(x, y, z+l) - f(x, y, z)].$$

By the mean value theorem for a function of one variable these differences may be expressed in the form

$$hf_x(x+\theta_1 h,\ y+k,\ z+l),\quad kf_y(x,\ y+\theta_2 k,\ z+l),\quad lf_z(x,\ y,\ z+\theta_3 l),$$

where θ_1, θ_2, θ_3 all lie between 0 and 1. Hence

$$\delta f = hf_x(x+\theta_1 h,\ y+k,\ z+l) + kf_y(x,\ y+\theta_2 k,\ z+l)$$
$$+ lf_z(x,\ y,\ z+\theta_3 l), \quad \text{......(1)}$$

or, since these derivatives are continuous and therefore tend to f_x, f_y, f_z respectively when h, k, l all tend to zero

$$\delta f = h(f_x+\omega_1) + k(f_y+\omega_2) + l(f_z+\omega_3)$$
$$= hf_x + kf_y + lf_z + (h\omega_1 + k\omega_2 + l\omega_3) \quad \text{......................(2)}$$

where ω_1, ω_2, ω_3 tend to zero when h, k, l all tend to zero.

Suppose now that x, y and z are functions of other variables, say functions of the two independent variables s and t, and that these functions and their partial derivatives with respect to s and t are continuous; the function $f(x, y, z)$ is now a function of functions and its derivatives $\partial f/\partial s$ and $\partial f/\partial t$ may be obtained at once by applying (2).

Let s alone vary and let the increments h or δx, k or δy, l or δz and δf correspond to the increment δs of s; if each member of (2) is divided by δs and the limit taken for $\delta s \to 0$ the last three terms in (2) will tend to zero because ω_1, ω_2, ω_3 tend to zero while their coefficients are finite. We thus find $\partial f/\partial s$ and by a similar process $\partial f/\partial t$, their expressions being as follows:

$$\left.\begin{aligned}\frac{\partial f}{\partial s} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s},\\ \frac{\partial f}{\partial t} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}.\end{aligned}\right\} \quad \text{....................(3)}$$

The method is obviously the same whatever be the number of variables in (2) or in (3); if x, y and z were functions of one variable only, say functions of t, the notation dx/dt, ... instead of $\partial x/\partial t$ ... would be used.

The equation (1) can be expressed so that instead of three fractions θ_1, θ_2, θ_3 there shall be only one, and this alternative form is often useful; it has in fact been already given in § 157,

p. 409, of the *Elementary Treatise* for the case of two variables. Let $F(t)=f(x+ht,\ y+kt,\ z+lt)$; then by the mean value theorem for a function of t

$$F(t)-F(0)=tF'(\theta t) \qquad 0<\theta<1.$$

It is proved in § 157 that

$$F'(t)=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}+l\frac{\partial}{\partial z}\right)f(x+ht,\ y+kt,\ z+lt),$$

and if we now put θt for t in this expression for $F'(t)$ and make t equal to unity we find

$$f(x+h,\ y+k,\ z+l)-f(x,\ y,\ z)=h\bar{f}_x+k\bar{f}_y+l\bar{f}_z, \qquad \text{.........(4)}$$

where $\bar{f}_x$ means the value of $\partial f(x,\ y,\ z)/\partial x$ when $x+\theta h$, $y+\theta k$ and $z+\theta l$ have been substituted in it for x, y and z respectively with similar meanings for $\bar{f}_y$, $\bar{f}_z$. This notation is not at all suggestive; an alternative notation is $f_x(x+\theta h,\ y+\theta k,\ z+\theta l)$ which is suggestive but cumbrous.

The equation (4) gives the Mean Value Theorem for functions of more than one variable.

Note. The earlier examples in the *Exercises* at the end of the chapter should be worked at this stage.

46. Differentials. The equation (2) of the last article has an important meaning in itself. For many purposes a valid approximation for the increment δf is desirable, and such an approximation is deducible from (2); but the expression for the approximation has also very useful applications to the problem of differentiation and change of variables, so that it is very desirable that the student should have a thorough grasp of the meaning and working of the differential, as the expression for the above approximation is called.

Let a principal infinitesimal (*E.T.* p. 195), ϱ say, be chosen when δx, δy, δz or h, k, l are infinitesimals; for example, let

$$\varrho=\sqrt{\{(\delta x)^2+(\delta y)^2+(\delta z)^2\}} \ \textit{ or } \ \varrho=|\,\delta x\,|+|\,\delta y\,|+|\,\delta z\,|$$

so that h/ϱ, k/ϱ and l/ϱ can not exceed unity (numerically). The part $h\omega_1+k\omega_2+l\omega_3$ of the increment of δf is an infinitesimal of a higher order than ϱ since its ratio to ϱ is numerically less than the sum

$$|\,\omega_1\,|+|\,\omega_2\,|+|\,\omega_3\,|$$

which tends to zero when h, k, l tend to zero. The other part

of the increment δf is called the *differential* of the function $f(x, y, z)$ and is denoted by $df(x, y, z)$ or df, so that

$$df = f_x \delta x + f_y \delta y + f_z \delta z. \qquad \text{.....(1)}$$

The differential df is a valid approximation for the increment δf when each of the increments δx, δy, δz is "small"; this approximation is of very frequent application.

There is, however, another aspect in which the subject may be considered, namely, that in which the differential sums up a whole set of derivatives; there is no question now of approximations but rather of useful rules of operation.

If x, y, z are *independent variables* the part

$$f_x \delta x + f_y \delta y + f_z \delta z, \text{ that is, } \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial z}\delta z,$$

of the increment δf is called the differential of $f(x, y, z)$ and is denoted, as before, by $df(x, y, z)$ or df simply. The term $f_x \delta x$ is a partial differential and may be denoted by $(df)_x$, with similar meanings for $(df)_y$ and $(df)_z$. The equation (1) is now considered simply as defining the differential *when the variables x, y, z are independent.*

Suppose now that x, y, z are not independent but are functions, say, of two independent variables s, t; the function $f(x, y, z)$ is therefore a function, say $F(s, t)$, of the independent variables s, t. By definition

$$df(x, y, z) = dF(s, t) = F_s \delta s + F_t \delta t. \qquad \text{.....(2)}$$

Now
$$F_s = f_x \frac{\partial x}{\partial s} + f_y \frac{\partial y}{\partial s} + f_z \frac{\partial z}{\partial s},$$

$$F_t = f_x \frac{\partial x}{\partial t} + f_y \frac{\partial y}{\partial t} + f_z \frac{\partial z}{\partial t},$$

while, since x, y, z are functions of the independent variables s, t, their differentials are given by the equations

$$dx = \frac{\partial x}{\partial s}\delta s + \frac{\partial x}{\partial t}\delta t, \quad dy = \ldots, \; dz = \ldots.$$

Hence

$$df = dF(s, t) = f_x\left(\frac{\partial x}{\partial s}\delta s + \frac{\partial x}{\partial t}\delta t\right) + f_y(\ldots) + f_z(\ldots),$$

so that
$$df = f_x dx + f_y dy + f_z dz. \qquad \text{.....(3)}$$

The forms (1) and (3) for the differential df differ in notation; in (3) the differentials dx, dy, dz take the place of the

increments δx, δy, δz respectively. Now the increments are arbitrary and no confusion can arise if they be expressed in the notation of differentials (see § 35); in fact, if a function of a single variable $f(x)$ be taken as a special case of a function $f(x, y, z)$ of three variables we can take x itself as the function and then $dx = f_x \delta x = \delta x$. With this change of notation we now have the theorem

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz, \quad \text{.................(4)}$$

whether the variables x, y, z are independent or not; df is *the total differential* of the function $f(x, y, z)$.

Note. If by any process the total differential df of a function of any number of *independent* variables s, t, u, ... has been expressed in the form

$$df = P\,ds + Q\,dt + R\,du + \ldots \quad \text{..................(5)}$$

where the differentials $ds, dt, du, \ldots$ of the independent variables alone appear and $P, Q, R, \ldots$ do not contain the differentials, then

$$P = \frac{\partial f}{\partial s}, \quad Q = \frac{\partial f}{\partial t}, \quad R = \frac{\partial f}{\partial u}, \ldots$$

For, by definition,

$$df = \frac{\partial f}{\partial s} ds + \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial u} du + \ldots \quad \text{(5')}$$

and the expressions (5) and (5′) for df must be identical; since the variables are independent we may suppose every increment except one, say ds, to be zero, and then we find $P = \partial f/\partial s$. Similar reasoning leads to the stated values of Q, R, ... We thus see that an equation such as (4) or (5) sums up a whole set of partial derivatives.

Ex. 1. If $u = (x^2 - y^2)/(x^2 + y^2)$ and $z = \sin^{-1}(u^{\frac{1}{2}})$ find dz.

$$dz = \frac{1}{(1-u)^{\frac{1}{2}}} \cdot \frac{1}{2} u^{-\frac{1}{2}} du = \frac{1}{2^{\frac{1}{2}}} \cdot \frac{x^2 + y^2}{y(x^2 - y^2)^{\frac{1}{2}}} du;$$

$$du = \frac{(x^2 + y^2)\,d(x^2 - y^2) - (x^2 - y^2)\,d(x^2 + y^2)}{(x^2 + y^2)^2} = \frac{4xy(y\,dx - x\,dy)}{(x^2 + y^2)^2},$$

and therefore

$$dz = \frac{2^{\frac{1}{2}} x(y\,dx - x\,dy)}{(x^2 + y^2)(x^2 - y^2)^{\frac{1}{2}}}.$$

The coefficients of dx and dy are $\partial z/\partial x$ and $\partial z/\partial y$ respectively, and may therefore be written down at once.

Ex. 2. If z is given as a function of two independent variables x and y, change the variables so that x becomes the function and z and y the independent variables, and express $\partial x/\partial z$ and $\partial x/\partial y$ in terms of the derivatives of z with respect to x and y.

When x and y are independent variables and z the dependent, a usual notation (which will be often employed) is

$$\frac{\partial z}{\partial x}=p, \quad \frac{\partial z}{\partial y}=q, \quad \frac{\partial^2 z}{\partial x^2}=r, \quad \frac{\partial^2 z}{\partial x \partial y}=s, \quad \frac{\partial^2 z}{\partial y^2}=t.$$

First, express dz in terms of dx and dy; we have

$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy=p\,dx+q\,dy. \quad \text{(i)}$$

Next, when z and y are independent and x the function, we have

$$dx=\frac{\partial x}{\partial z}dz+\frac{\partial x}{\partial y}dy. \quad \text{(ii)}$$

Again, by solving (i) for dx we find

$$dx=\frac{1}{p}dz-\frac{q}{p}dy. \quad \text{(iii)}$$

Equations (ii) and (iii) must be identical by the above *Note*, and therefore

$$\frac{\partial x}{\partial z}=\frac{1}{p}, \quad \frac{\partial x}{\partial y}=-\frac{q}{p}.$$

See example of § 47 (*Second Method*) for a different treatment.

Ex. 3. If u and v are determined as functions of x, y, z by the equations $\varphi(x, y, z, u, v)=0$ and $\psi(x, y, z, u, v)=0$, find the derivatives of u and v with respect to x, y, z.

In the solution, to secure brevity, we use the notation of Jacobians (§ 55), namely,

$$\frac{\partial(\varphi, \psi)}{\partial(u, v)}=\begin{vmatrix}\varphi_u, & \varphi_v\\ \psi_u, & \psi_v\end{vmatrix}=\varphi_u\psi_v-\varphi_v\psi_u\,; \quad \frac{\partial(\varphi, \psi)}{\partial(v, x)}=\begin{vmatrix}\varphi_v, & \varphi_x\\ \psi_v, & \psi_x\end{vmatrix}$$

and so on.

First Method. Differentiate $\varphi=0$ and $\psi=0$ with respect to x, keeping y and z constant; thus

$$\varphi_x+\varphi_u\frac{\partial u}{\partial x}+\varphi_v\frac{\partial v}{\partial x}=0, \quad \psi_x+\psi_u\frac{\partial u}{\partial x}+\psi_v\frac{\partial v}{\partial x}=0.$$

If these equations be solved for $\partial u/\partial x$ and $\partial v/\partial x$ we get

$$\frac{\partial u}{\partial x}=\frac{\partial(\varphi, \psi)}{\partial(v, x)}\div\frac{\partial(\varphi, \psi)}{\partial(u, v)}, \quad \frac{\partial v}{\partial x}=\frac{\partial(\varphi, \psi)}{\partial(x, u)}\div\frac{\partial(\varphi, \psi)}{\partial(u, v)}.$$

In the same way the other derivatives may be found.

Second Method. Use differentials. The total differentials $d\varphi$ and $d\psi$ are zero; therefore

$$\varphi_x\,dx+\varphi_y\,dy+\varphi_z\,dz+\varphi_u\,du+\varphi_v\,dv=0,$$
$$\psi_x\,dx+\psi_y\,dy+\psi_z\,dz+\psi_u\,du+\psi_v\,dv=0.$$

Solve these equations for du and dv; we get

$$\frac{\partial(\varphi,\psi)}{\partial(u,v)}du=\frac{\partial(\varphi,\psi)}{\partial(v,x)}dx+\frac{\partial(\varphi,\psi)}{\partial(v,y)}dy+\frac{\partial(\varphi,\psi)}{\partial(v,z)}dz,$$

$$\frac{\partial(\varphi,\psi)}{\partial(u,v)}dv=\frac{\partial(\varphi,\psi)}{\partial(x,u)}dx+\frac{\partial(\varphi,\psi)}{\partial(y,u)}dy+\frac{\partial(\varphi,\psi)}{\partial(z,u)}dz.$$

But
$$du=\frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy+\frac{\partial u}{\partial z}dz;\ dv=\frac{\partial v}{\partial x}dx+\dots.$$

The equating of coefficients of dx, dy, dz gives

$$\frac{\partial u}{\partial x}=\frac{\partial(\varphi,\psi)}{\partial(v,x)}\div\frac{\partial(\varphi,\psi)}{\partial(u,v)},$$

and the other derivatives may similarly be written down at once.

47. Higher Differentials. For brevity suppose that f is a function of two variables x and y. The second differential of $f(x, y)$ is the differential of $df(x, y)$, that is, $d[df(x, y)]$ which is denoted by $d^2f(x, y)$ or simply d^2f. Hence

$$d^2f=d(f_x dx+f_y dy)=dx\,df_x+f_x d(dx)+dy\,df_y+f_y d(dy),$$

or
$$d^2f=dx\,df_x+dy\,df_y+f_x d^2x+f_y d^2y. \quad\dots\dots\dots\dots(1)$$

Now
$$df_x=\frac{\partial f_x}{\partial x}dx+\frac{\partial f_x}{\partial y}dy=f_{xx}dx+f_{xy}dy,$$

and
$$df_y=\frac{\partial f_y}{\partial x}dx+\frac{\partial f_y}{\partial y}dy=f_{xy}dx+f_{yy}dy,$$

so that

$$\begin{aligned}dx\,df_x+dy\,df_y&=f_{xx}(dx)^2+2f_{xy}dx\,dy+f_{yy}(dy)^2\\&=f_{xx}dx^2+2f_{xy}dx\,dy+f_{yy}dy^2,\end{aligned}$$

and therefore

$$d^2f=f_{xx}dx^2+2f_{xy}dx\,dy+f_{yy}dy^2+f_x d^2x+f_y d^2y. \quad \dots \quad (2)$$

The question now arises "how is the distinction made between the case of x, y as independent variables and that of x, y as functions of other variables"? The answer that is found to be most convenient is that, when x and y are independent variables their differentials dx and dy are taken to be constant so that d^2x and d^2y are zero. Hence when x and y are *independent*

$$d^2f(x,y)=d^2f=f_{xx}dx^2+2f_{xy}dx\,dy+f_{yy}dy^2, \quad\dots\dots\dots(3)$$

while if x and y are not independent $d^2f(x, y)$ is given by (2).

In the same way the third differential $d(d^2f)$ or d^3f and higher differentials are defined; when x and y are independent

$d^3x, d^4x, \ldots$ and $d^3y, d^4y, \ldots$ are all zero. Thus one part of d^3f is formed from the first term in the expression (3) for d^2f and gives

$$d[f_{xx}dx^2] = dx^2\, df_{xx} = dx^2\left[\frac{\partial f_{xx}}{\partial x}\,dx + \frac{\partial f_{xx}}{\partial y}\,dy\right],$$

or $$f_{xxx}dx^3 + f_{xxy}dx^2\,dy, \qquad\ldots\ldots(4)$$

and for the value of d^3f when x and y are independent we find

$$d^3f = f_{xxx}dx^3 + 3f_{xxy}\,dx^2dy + 3f_{xyy}\,dx\,dy^2 + f_{yyy}\,dy^3. \qquad\ldots\ldots(5)$$

If x and y are not independent we must in finding d^3f take equation (2) and include the terms that arise from dx and dy which are no longer constant. Thus to find $d[f_{xx}dx^2]$ we must add to the expression (4) the term $f_{xx}d(dx^2)$ or $2f_{xx}dx\,d^2x$ since $d(dx^2)$ is $2dx\,d^2x$.

The notation of differentials gives a compact form for Taylor's expansion of $f(x+h, y+k)$ when $h = dx$, $k = dy$ and h, k are constant; the form is simply

$$f(x+h, y+k) = f + df + \frac{d^2f}{2!} + \frac{d^3f}{3!} + \ldots \qquad\ldots\ldots(6)$$

and the series on the right is of the same form whatever be the number of *independent* variables $x, y, z, \ldots$, it being noted that

$$df(x, y, z, \ldots) = f_x dx + f_y dy + f_z dz + \cdots.$$

(See *E.T.* p. 508.)

It must be specially noted, however, that when $x, y, z, \ldots$ are not independent the expressions for $d^2f, d^3f, \ldots$ have no longer the simple forms given by (3) or (5) but involve the higher differentials of $dx, dy, dz, \ldots$.

Ex. Take Example 2, § 46, and express the second derivatives of x with respect to z and y in terms of p, q, r, s, t.

First Method. We have $\dfrac{\partial x}{\partial z} = \dfrac{1}{p}$, $\dfrac{\partial x}{\partial y} = -\dfrac{q}{p}$. $\qquad\ldots\ldots$(i)

From the first of these equations we find by taking the differentials

$$d\left(\frac{\partial x}{\partial z}\right) = d\left(\frac{1}{p}\right), \text{ or, } \frac{\partial^2 x}{\partial z^2}dz + \frac{\partial^2 x}{\partial z\partial y}dy = -\frac{dp}{p^2}. \qquad\ldots\ldots\text{(ii)}$$

Now $$dp = \frac{\partial p}{\partial x}dx + \frac{\partial p}{\partial y}dy = r\,dx + s\,dy,$$

since p is given as a function of x and y. In this expression for dp put the value of dx in terms of dz and dy, namely $(dz - q\,dy)/p$, and the equation (ii) takes the form

$$\frac{\partial^2 x}{\partial z^2}dz + \frac{\partial^2 x}{\partial z\partial y}dy = -\frac{r}{p^3}dz + \frac{rq - sp}{p^3}dy.$$

In this equation we have only the differentials of independent variables, and can therefore equate the coefficients of dz and dy respectively; hence

$$\frac{\partial^2 x}{\partial z^2} = -\frac{r}{p^3}, \quad \frac{\partial^2 x}{\partial z \partial y} = \frac{rq - sp}{p^3}.$$

In the same way the second of equations (i) gives

$$\frac{\partial^2 x}{\partial z \partial y} dz + \frac{\partial^2 x}{\partial y^2} dy = -\frac{p\,dq - q\,dp}{p^2}.$$

But $$dq = \frac{\partial q}{\partial x} dx + \frac{\partial q}{\partial y} dy = s\,dx + t\,dy,$$

and substituting as before for dx in dq and dp we find

$$\frac{\partial^2 x}{\partial z \partial y} dz + \frac{\partial^2 x}{\partial y^2} dy = \frac{rq - sp}{p^3} dz + \frac{2pqs - tp^2 - rq^2}{p^3} dy,$$

and therefore $$\frac{\partial^2 x}{\partial y^2} = \frac{2pqs - tp^2 - rq^2}{p^3},$$

the value of $\partial^2 x/\partial z \partial y$ being the same as before.

Cor. $$\frac{\partial^2 x}{\partial z^2} \cdot \frac{\partial^2 x}{\partial y^2} - \left(\frac{\partial^2 x}{\partial z \partial y}\right)^2 = \frac{rt - s^2}{p^4}.$$

Second Method. Suppose $z = f(x, y)$; then

$$p = \frac{\partial z}{\partial x} = f_x, \quad q = \frac{\partial z}{\partial y} = f_y.$$

If the equation $z = f(x, y)$ is solved for x in terms of z and y, giving $x = \varphi(z, y)$ say, then the partial derivatives with respect to z and y of any function $\psi(x, y)$ are given by

$$\left(\frac{\partial \psi}{\partial z}\right) = \frac{\partial \psi}{\partial x} \frac{\partial x}{\partial z}, \quad \left(\frac{\partial \psi}{\partial y}\right) = \frac{\partial \psi}{\partial x} \frac{\partial x}{\partial y} + \frac{\partial \psi}{\partial y}.$$

Here $\left(\frac{\partial \psi}{\partial z}\right)$ and $\left(\frac{\partial \psi}{\partial y}\right)$ are the *total* derivatives of ψ with respect to z and y. The function $\psi(x, y)$ contains z, since x is now a function of z and y, and contains y *explicitly* as well as through x, thus bringing in the additional term $\partial \psi/\partial y$. The two values $(\partial \psi/\partial y)$ and $\partial \psi/\partial y$ are quite different.

Now differentiate the equation $z = f(x, y)$ with respect to z and y respectively; we find

$$\frac{\partial z}{\partial z} = \left(\frac{\partial f(x, y)}{\partial z}\right), \text{ that is, } 1 = f_x \frac{\partial x}{\partial z}, \text{ so that } \frac{\partial x}{\partial z} = \frac{1}{f_x} = \frac{1}{p};$$

$$0 = \left(\frac{\partial f(x, y)}{\partial y}\right), \text{ that is, } 0 = f_x \frac{\partial x}{\partial y} + f_y, \text{ so that } \frac{\partial x}{\partial y} = -\frac{q}{p}.$$

Again, $$\frac{\partial^2 x}{\partial z^2} = \frac{-1}{p^2}\left(\frac{\partial p}{\partial z}\right) = \frac{-1}{p^2} \cdot \frac{\partial p}{\partial x} \frac{\partial x}{\partial z} = \frac{-r}{p^3},$$

since p is a function of x and y.

$$\frac{\partial^2 x}{\partial z\,\partial y}=\frac{-1}{p^2}\left(\frac{\partial p}{\partial y}\right)=\frac{-1}{p^2}\left[\frac{\partial p}{\partial x}\frac{\partial x}{\partial y}+\frac{\partial p}{\partial y}\right]=\frac{rq-sp}{p^3}.$$

$$\frac{\partial^2 x}{\partial y^2}=\frac{-1}{p^2}\left[p\left(\frac{\partial q}{\partial y}\right)-q\left(\frac{\partial p}{\partial y}\right)\right]=\frac{2pqs-tp^2-rq^2}{p^3},$$

since
$$\left(\frac{\partial q}{\partial y}\right)=\frac{\partial q}{\partial x}\frac{\partial x}{\partial y}+\frac{\partial q}{\partial y}=-s\frac{q}{p}+t.$$

The student should note carefully the two different meanings of the symbol $\partial\psi(x, y)/\partial y$. When x is independent of y there is no ambiguity, the derivative being as usual; but if x contains y the derivative is a *total* derivative (see *E.T.* p. 212 and p. 219).

If the relation between x, y and z is given by an equation of the form $F(x, y, z)=0$ we have, when z is considered as a function of x and y,

$$F_x+F_z\frac{\partial z}{\partial x}=0, \quad \text{or} \quad p=-F_x/F_z,$$

and similarly $q=-F_y/F_z$.

When $F=0$ is taken as defining x as a function of z and y we find

$$F_z+F_x\frac{\partial x}{\partial z}=0, \quad \frac{\partial x}{\partial z}=-\frac{F_z}{F_x}=\frac{1}{p};$$

$$F_y+F_x\frac{\partial x}{\partial y}=0, \quad \frac{\partial x}{\partial y}=-\frac{F_y}{F_x}=-\frac{q}{p}.$$

The rest of the work is as before.

48. Change of Variables. If y is given as a function of x, and if the variables are changed by the substitutions $x=\varphi(u, v)$, $y=\psi(u, v)$ so that v becomes a function of u, the problem of expressing the derivatives of y with respect to x in terms of u, v and the derivatives of v with respect to u has been briefly discussed (*E.T.* p. 234) for the simple but important case $x=v\cos u$, $y=v\sin u$. The method is quite general and the student should have little difficulty in applying it to a given case. Thus

$$\frac{dy}{dx}=\frac{dy}{du}\div\frac{dx}{du}=\left(\psi_u+\psi_v\frac{dv}{du}\right)\div\left(\varphi_u+\varphi_v\frac{dv}{du}\right),$$

$$\frac{d^2y}{dx^2}=\left(\frac{d}{du}\cdot\frac{dy}{dx}\right)\div\frac{dx}{du},$$

and the derivatives $\dfrac{d}{du}\left(\dfrac{dy}{dx}\right)$ and $\dfrac{dx}{du}$ are easily found.

The problem of change of variables for functions of several variables is distinctly harder. A special but very important case has been worked out fully (*E.T.* pp. 237-240), and the principles that underlie the solution in that case will now be

illustrated in some detail, particularly for a function of two independent variables.

Problem I a. *If z is a function $f(x, y)$ of the independent variables x, y and if x, y are changed to new independent variables u, v by the substitutions $x=\varphi(u, v)$, $y=\psi(u, v)$ it is required to express the derivatives of z with respect to x, y in terms of u, v and the derivatives of z with respect to u, v.*

The student must pay particular attention to the meaning of the symbols. Thus $\partial z/\partial x$ means the partial derivative of z with respect to x when y is constant, while $\partial z/\partial u$ is the partial derivative of z with respect to u when v is constant, *z being now expressed in terms of u, v.*

By § 45, (3), we find

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial x}, \quad \frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial y}. \quad \text{.........(1)}$$

To obtain $\partial u/\partial x, \ldots, \partial v/\partial y$, differentiate the equations $x=\varphi(u, v)$ and $y=\psi(u, v)$ with respect to x and y respectively; thus, differentiating with respect to x, we get

$$1=\frac{\partial \varphi}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial \varphi}{\partial v}\frac{\partial v}{\partial x}, \quad \text{or,} \quad 1=\varphi_u\frac{\partial u}{\partial x}+\varphi_v\frac{\partial v}{\partial x},$$

$$0=\frac{\partial \psi}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial \psi}{\partial v}\frac{\partial v}{\partial x}, \quad \text{or,} \quad 0=\psi_u\frac{\partial u}{\partial x}+\psi_v\frac{\partial v}{\partial x},$$

and these equations give

$$\frac{\partial u}{\partial x}=\frac{\psi_v}{J}, \quad \frac{\partial v}{\partial x}=-\frac{\psi_u}{J} \quad \text{.....................(2a)}$$

while, by differentiating with respect to y, we find

$$\frac{\partial u}{\partial y}=-\frac{\varphi_v}{J}, \quad \frac{\partial v}{\partial y}=\frac{\varphi_u}{J} \quad \text{.....................(2b)}$$

where J is the Jacobian (see § 46, Ex. 3 or § 55)

$$J=\varphi_u\psi_v-\varphi_v\psi_u=\frac{\partial(\varphi, \psi)}{\partial(u, v)}=\frac{\partial(x, y)}{\partial(u, v)}.$$

The equations (1) now give the required values of $\partial z/\partial x$ and $\partial z/\partial y$, namely,

$$\frac{\partial z}{\partial x}=\frac{\psi_v}{J}\frac{\partial z}{\partial u}-\frac{\psi_u}{J}\frac{\partial z}{\partial v}, \quad \frac{\partial z}{\partial y}=-\frac{\varphi_v}{J}\frac{\partial z}{\partial u}+\frac{\varphi_u}{J}\frac{\partial z}{\partial v}. \quad \text{.........(3)}$$

We may also proceed as follows. By § 45, (3) we have

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \quad \frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v}, \quad (1')$$

and if these equations be solved for $\partial z/\partial x$ and $\partial z/\partial y$ in terms of $\partial z/\partial u$

and $\partial z/\partial v$ we find the values given by equations (3), it being observed that $\partial x/\partial u, \ldots, \partial y/\partial v$ mean the same thing as $\varphi_u, \ldots, \psi_v$ respectively.

Equations (3) are of the form

$$\frac{\partial z}{\partial x}=A\frac{\partial z}{\partial u}+B\frac{\partial z}{\partial v},\quad \frac{\partial z}{\partial y}=C\frac{\partial z}{\partial u}+D\frac{\partial z}{\partial v}. \qquad\ldots\ldots\ldots\ldots\ldots(4)$$

where A, B, C, D, $\partial z/\partial u$, $\partial z/dv$ are all functions of u, v and do not contain x, y explicitly. We may therefore say that $\partial z/\partial x$ and $\partial z/\partial y$ are functions, $F(u, v)$ and $G(u, v)$ say, of the new variables u, v. Hence, to find $\partial^2 z/\partial x^2$ we put $F(u, v)$ in place of z in the first of equations (3) or (4); thus

$$\frac{\partial^2 z}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right)=\frac{\partial}{\partial x}\,.\,F(u, v)=A\frac{\partial F}{\partial u}+B\frac{\partial F}{\partial v}. \qquad\ldots\ldots\ldots\ldots(5)$$

the function $\partial z/\partial x$ of x and y being $F(u, v)$ when expressed in terms of u and v. In the same way we see that

$$\frac{\partial^2 z}{\partial y^2}=\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right)=\frac{\partial}{\partial y}\,.\,G(u, v)=C\frac{\partial G}{\partial u}+D\frac{\partial G}{\partial v}$$

$$\frac{\partial^2 z}{\partial x\partial y}=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)=\frac{\partial}{\partial x}\,.\,G(u, v)=A\frac{\partial G}{\partial u}+B\frac{\partial G}{\partial v}. \qquad\ldots\ldots\ldots(5a)$$

or

$$\frac{\partial^2 z}{\partial x\partial y}=\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)=\frac{\partial}{\partial y}\,.\,F(u, v)=C\frac{\partial F}{\partial u}+D\frac{\partial F}{\partial v}$$

When the derivatives are expressed as in (4) we may say that the operators $\partial/\partial x$ and $\partial/\partial y$, applied to any function w of x and y, are equivalent to the operators

$$A\frac{\partial}{\partial u}+B\frac{\partial}{\partial v} \quad\text{and}\quad C\frac{\partial}{\partial u}+D\frac{\partial}{\partial v},$$

respectively, acting on w_1 where w_1 is the value of w expressed in terms of u and v.

The value of $\partial^2 z/\partial x^2$ is thus seen to be

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2}&=A\left(A\frac{\partial^2 z}{\partial u^2}+B\frac{\partial^2 z}{\partial u\,\partial v}+\frac{\partial A}{\partial u}\frac{\partial z}{\partial u}+\frac{\partial B}{\partial u}\frac{\partial z}{\partial v}\right)\\&\quad+B\left(A\frac{\partial^2 z}{\partial u\,\partial v}+B\frac{\partial^2 z}{\partial v^2}+\frac{\partial A}{\partial v}\frac{\partial z}{\partial u}+\frac{\partial B}{\partial v}\frac{\partial z}{\partial v}\right)\\&=A^2\frac{\partial^2 z}{\partial u^2}+2AB\frac{\partial^2 z}{\partial u\partial v}+B^2\frac{\partial^2 z}{\partial v^2}+\left(A\frac{\partial A}{\partial u}+B\frac{\partial A}{\partial v}\right)\frac{\partial z}{\partial u}\\&\qquad+\left(A\frac{\partial B}{\partial u}+B\frac{\partial B}{\partial v}\right)\frac{\partial z}{\partial v},\end{aligned}$$

and the values of $\partial^2 z/\partial y^2$ and $\partial^2 z/\partial x\,\partial y$ may be found in the same way.

The higher derivatives may be obtained by exactly the same method; fortunately they are not often required. The algebra of the transformation is tedious but the method seems simple.

Ex. 1. If $x=r\cos\theta$, $y=r\sin\theta$ show that

$$\frac{\partial^2 z}{\partial x\,\partial y}=\cos\theta\sin\theta\frac{\partial^2 z}{\partial r^2}+\frac{\cos^2\theta-\sin^2\theta}{r}\frac{\partial^2 z}{\partial r\,\partial\theta}-\frac{\cos\theta\sin\theta}{r^2}\frac{\partial^2 z}{\partial\theta^2}$$
$$-\frac{\cos\theta\sin\theta}{r}\frac{\partial z}{\partial r}-\frac{\cos^2\theta-\sin^2\theta}{r^2}\frac{\partial z}{\partial\theta}.$$

See *E.T.* p. 236, equations (3) and (4).

Problem I b. *If the relation between the old and the new variables is given by the equations* $u=\varphi(x, y)$, $v=\psi(x, y)$, *express the old derivatives in terms of the new.*

It is of course understood that the equations determine x and y as functions of u and v, or that they can be solved for x and y in terms of u and v. In this case the form of the solution is simple; we have

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\varphi_x+\frac{\partial z}{\partial v}\psi_x,\quad \frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\varphi_y+\frac{\partial z}{\partial v}\psi_y,$$

where the derivatives φ_x, ψ_x, φ_y, ψ_y must now be expressed in terms of u and v. The higher derivatives are then found as before.

Problem I c. *If the relation between the old and the new variables is given by the equations* $\varphi(x, y, u, v)=0$, $\psi(x, y, u, v)=0$, *express the old derivatives in terms of the new.*

Assuming that the given equations define each pair of the variables in terms of the other pair we proceed as follows. The functions to be determined are the functions A, B, C, D of equations (4). It may be possible in a given case to find the expressions for u, v in terms of x, y or for x, y in terms of u, v, and when these are found we can apply one of the two methods already given. If the expressions just mentioned cannot be found conveniently we calculate $\partial u/\partial x, \dots, \partial v/\partial y$ as in § 46, Ex. 3; these values are

$$\frac{\partial u}{\partial x}=\frac{\partial(\varphi, \psi)}{\partial(v, x)}\div\frac{\partial(\varphi, \psi)}{\partial(u, v)},\quad \frac{\partial v}{\partial x}=\frac{\partial(\varphi, \psi)}{\partial(x, u)}\div\frac{\partial(\varphi, \psi)}{\partial(u, v)}.$$
$$\dots\dots\dots(6)$$
$$\frac{\partial u}{\partial y}=\frac{\partial(\varphi, \psi)}{\partial(v, y)}\div\frac{\partial(\varphi, \psi)}{\partial(u, v)},\quad \frac{\partial v}{\partial y}=\frac{\partial(\varphi, \psi)}{\partial(y, u)}\div\frac{\partial(\varphi, \psi)}{\partial(u, v)}.$$

When these values of the derivatives of u and v are inserted in the equations (3) we have the formal expressions for $\partial z/\partial x$ and $\partial z/\partial y$; if in these expressions the values of x and y in terms of u and v are substituted we shall have equations (4) and can then find the higher derivatives as before. Even when x, y cannot be conveniently expressed in terms of u, v the values of $\partial u/\partial x, \ldots, \partial v/\partial y$ given by equations (6) are often useful, and the student should note carefully the method by which they are obtained.

Ex. 2. Apply equations (6) to show that

$$\text{(i)}\ \frac{\partial(\varphi, \psi)}{\partial(u, v)} \cdot \frac{\partial(u, v)}{\partial(x, y)} = \frac{\partial(\varphi, \psi)}{\partial(x, y)}; \quad \text{(ii)}\ \frac{\partial(x, y)}{\partial(u, v)} \cdot \frac{\partial(u, v)}{\partial(x, y)} = 1.$$

Problem II. *If z is a function $F(x, y)$ of the independent variables x, y, and if all the variables are changed by the substitution*

$$x = f(u, v, w), \quad y = g(u, v, w), \quad z = h(u, v, w), \quad \ldots\ldots\ldots(7)$$

it is required to express $\partial z/\partial x$ and $\partial z/\partial y$ in terms of u, v and the derivatives $\partial w/\partial u$ and $\partial w/\partial v$ of the new function w with respect to the new independent variables u, v.

It is supposed that equations (7) determine u, v, w as functions of x, y, z. The equation $z = F(x, y)$ becomes an equation between u, v, w which defines a function w of u, v. Hence z may be considered as a function of x, y where x, y are functions of u, v, w, and w a function of u, v; we may therefore find $(\partial z/\partial u)$ and $(\partial z/\partial v)$ in terms of $\partial z/\partial x$ and $\partial z/\partial y$ by the rule for "function of a function." The forms in brackets are meant to indicate that z is a function of u, v and another variable w which is also a function of u, v.

Thus when $z = h(u, v, w)$

$$\left(\frac{\partial z}{\partial u}\right) = \frac{\partial z}{\partial u} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial u} = h_u + h_w \frac{\partial w}{\partial u} \quad \ldots\ldots\ldots\ldots\ldots\ldots(8)$$

and there are analogous expressions for $(\partial x/\partial u), \ldots, (\partial y/\partial v)$.

Now $$\left(\frac{\partial z}{\partial u}\right) = \frac{\partial z}{\partial x}\left(\frac{\partial x}{\partial u}\right) + \frac{\partial z}{\partial y}\left(\frac{\partial y}{\partial u}\right),$$

since $z = F(x, y)$ and x, y are functions of u, v, w. If we now insert the values of $(\partial z/\partial u)$, $(\partial x/\partial u)$ and $(\partial y/\partial u)$ as given

by (8) and apply the same method to $(\partial z/\partial v)$ we find the equations

$$\left.\begin{aligned} h_u+h_w\frac{\partial w}{\partial u}&=\left(f_u+f_w\frac{\partial w}{\partial u}\right)\frac{\partial z}{\partial x}+\left(g_u+g_w\frac{\partial w}{\partial u}\right)\frac{\partial z}{\partial y}\\ h_v+h_w\frac{\partial w}{\partial v}&=\left(f_v+f_w\frac{\partial w}{\partial v}\right)\frac{\partial z}{\partial x}+\left(g_v+g_w\frac{\partial w}{\partial v}\right)\frac{\partial z}{\partial y}\end{aligned}\right\}\quad\text{.........(9)}$$

Equations (9) when solved for $\partial z/\partial x$ and $\partial z/\partial y$ determine these derivatives in terms of $\partial w/\partial u$ and $\partial w/\partial v$. As a rule, the coefficients in the expressions will involve all the variables, but as x, y, z, w are all functions of u, v, the solution is theoretically complete though in practice the actual determination of the explicit forms in terms of u, v may be very laborious. The values given by (9) are, however, of great importance in many applications.

Equations (9) may be found by using differentials, thus :

$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy=p\,dx+q\,dy.$$

Now express dx, dy, dz in terms of du, dv ; then

$$\left(\frac{\partial z}{\partial u}\right)du+\left(\frac{\partial z}{\partial v}\right)dv=p\left[\left(\frac{\partial x}{\partial u}\right)du+\left(\frac{\partial x}{\partial v}\right)dv\right]+q\left[\left(\frac{\partial y}{\partial u}\right)du+\left(\frac{\partial y}{\partial v}\right)dv\right].$$

Equating coefficients of du and dv we find, using (8), the values given by (9).

49. Special Cases. In problems involving change of variables it is frequently required to transform a particular expression involving a combination of derivatives, and the general methods given above can often be modified so as to reduce the algebraic work. The following example illustrates an important case.

Ex. 1. Transform the expression $\dfrac{\partial^2 f}{\partial x^2}+\dfrac{\partial^2 f}{\partial y^2}$ by the substitution $x=\varphi(u, v)$, $y=\psi(u, v)$ and show that

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}=\frac{1}{J}\left(\frac{\partial^2 f}{\partial u^2}+\frac{\partial^2 f}{\partial v^2}\right),\ \dots (A),\ \text{where } J=\left(\frac{\partial x}{\partial u}\right)^2+\left(\frac{\partial x}{\partial v}\right)^2$$

if $$\frac{\partial x}{\partial u}=\frac{\partial y}{\partial v},\quad \frac{\partial x}{\partial v}=-\frac{\partial y}{\partial u}.\ \text{..............................(i)}$$

In this case it is much simpler to transform from derivatives in u, v to derivatives in x, y. No doubt, by this method the transformation is rather verified than proved, but the method of verification is important;

a straightforward application of the general method would be tedious. We have, using the subscript notation to save space,

$$f_u = f_x \frac{\partial x}{\partial u} + f_y \frac{\partial y}{\partial u}. \qquad \text{(ii)}$$

In finding f_{uu} the factors $\partial x/\partial u$ and $\partial y/\partial u$ are obtained by differentiating with respect to u; the values of $\partial f_x/\partial u$ and $\partial f_y/\partial u$ are found by substituting f_x and f_y respectively in place of f in (ii).

Thus $$f_{uu} = f_x \frac{\partial^2 x}{\partial u^2} + f_y \frac{\partial^2 y}{\partial u^2} + \frac{\partial x}{\partial u}\frac{\partial f_x}{\partial u} + \frac{\partial y}{\partial u}\frac{\partial f_y}{\partial u};$$

but, by (ii),

$$\frac{\partial f_x}{\partial u} = f_{xx}\frac{\partial x}{\partial u} + f_{xy}\frac{\partial y}{\partial u}, \quad \frac{\partial f_y}{\partial u} = f_{xy}\frac{\partial x}{\partial u} + f_{yy}\frac{\partial y}{\partial u},$$

so that $$f_{uu} = f_x \frac{\partial^2 x}{\partial u^2} + f_y \frac{\partial^2 y}{\partial u^2} + \left(\frac{\partial x}{\partial u}\right)^2 f_{xx} + 2\frac{\partial x}{\partial u}\frac{\partial y}{\partial u} f_{xy} + \left(\frac{\partial y}{\partial u}\right)^2 f_{yy}. \qquad \text{(iii)}$$

In the same way we find

$$f_{vv} = f_x \frac{\partial^2 x}{\partial v^2} + f_y \frac{\partial^2 y}{\partial v^2} + \left(\frac{\partial x}{\partial v}\right)^2 f_{xx} + 2\frac{\partial x}{\partial v}\frac{\partial y}{\partial v} f_{xy} + \left(\frac{\partial y}{\partial v}\right)^2 f_{yy}. \qquad \text{(iv)}$$

The expression $f_{uu} + f_{vv}$ involves some symmetrical combinations of the derivatives of x and y which in virtue of conditions (i) reduce the whole to a simple form. We have

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial x}{\partial v}\right)^2 = \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2. \qquad \text{(v)}$$

$$\frac{\partial^2 x}{\partial u^2} = \frac{\partial}{\partial u}\frac{\partial y}{\partial v} = \frac{\partial}{\partial v}\frac{\partial y}{\partial u} = -\frac{\partial^2 x}{dv^2}, \quad \frac{\partial^2 x}{\partial u^2} + \frac{\partial^2 x}{\partial v^2} = 0, \qquad \text{(vi)}$$

and in the same way $$\frac{\partial^2 y}{\partial u^2} + \frac{\partial^2 y}{\partial v^2} = 0. \qquad \text{(vii)}$$

Also $$\frac{\partial x}{\partial u}\frac{\partial y}{\partial u} + \frac{\partial x}{\partial v}\frac{\partial y}{\partial v} = -\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial x}{\partial v}\frac{\partial x}{\partial u} = 0. \qquad \text{(viii)}$$

If we now add corresponding sides of (iii) and (iv) and take account of (v) ... (viii) we find

$$f_{uu} + f_{vv} = J(f_{xx} + f_{yy}).$$

Hence, if f satisfies the equation $f_{xx} + f_{yy} = 0$ it also satisfies the equation $f_{uu} + f_{vv} = 0$ since $J \neq 0$.

If x and y are interpreted as rectangular coordinates the equations $x = \varphi(u, v)$, $y = \psi(u, v)$ determine, by assigning constant values to u and letting v vary, a family of curves (u = constant), while if v is constant and u variable the equations determine another family of curves (v = constant). The equation (viii) shows that the two families are orthogonal—that is, at each point where the two sets of curves intersect the tangent to the one curve is perpendicular to the tangent to the other. Whenever the curves are orthogonal the equation $f_{xx} + f_{yy} = 0$ becomes $f_{uu} + f_{vv} = 0$, that is, does not alter its form.

It is not hard to prove that, ***when the equations*** (i) ***are satisfied***, the equations (i), (vi), (vii) and (viii) will be satisfied if x and u are interchanged and also y and v—that is, ***if x and y become the independent variables and u and v functions of x and y***; the Jacobian J will become $1/J$. (See § 48, Problem I. a). The proof of the relation (A) by applying the general method may then be carried out, as above.

The student may, as an exercise, work out the transformation when $x=\cosh u\cos v$, $y=\sinh u\sin v$; the two families of curves are confocal ellipses and hyperbolas.

Ex. 2. If z is a function $f(x, y)$ of the independent variables x, y, and if the variables are changed to the independent variables u, v and the function w where

$$u=\frac{\partial z}{\partial x}=p,\quad v=\frac{\partial z}{\partial y}=q,\quad w=px+qy-z, \qquad\text{(i)}$$

find the first and second derivatives of w with respect to u, v.

Apply the method of differentials. We have

$$dw=p\,dx+x\,dp+q\,dy+y\,dq-dz=x\,dp+y\,dq$$

since
$$dz=p\,dx+q\,dy,$$

so that
$$dw=x\,dp+y\,dq=x\,du+y\,dv.$$

But
$$dw=\frac{\partial w}{\partial u}du+\frac{\partial w}{\partial v}dv,\quad\text{so that}\quad \frac{\partial w}{\partial u}=x,\quad \frac{\partial w}{\partial v}=y.$$

If $\partial w/\partial u=P$ and $\partial w/\partial v=Q$ we therefore have

$$x=\frac{\partial w}{\partial u}=P,\quad y=\frac{\partial w}{\partial v}=Q,\quad z=Pu+Qv-w. \qquad\text{(ii)}$$

Let $R=\partial^2 w/\partial u^2$, $S=\partial^2 w/\partial u\,\partial v$, $T=\partial^2 w/\partial v^2$, the symbols r, s, t denoting the corresponding derivatives of z as to x, y (§ 46, Ex. 2).

Take the differential of P; thus

$$dP=\frac{\partial P}{\partial u}du+\frac{\partial P}{\partial v}dv=R\,du+S\,dv.$$

Now, expressing these differentials in terms of dx and dy we get

$$dP=dx,\quad du=dp=r\,dx+s\,dy,\quad dv=dq=s\,dx+t\,dy,$$

so that
$$dx=R(r\,dx+s\,dy)+S(s\,dx+t\,dy),$$

and therefore, equating coefficients of dx and dy,

$$Rr+Ss=1,\quad Rs+St=0.$$

Hence
$$R=t/(rt-s^2),\quad S=-s/(rt-s^2),$$

and, by a similar method, $\quad T=r/(rt-s^2)$.

It is easily shown that $RT-S^2=1/(rt-s^2)$ and therefore

$$r=T/(RT-S^2),\ s=-S/(RT-S^2),\ t=R/(RT-S^2).$$

The transformation fails if $rt-s^2=0$.

The transformation is known as Legendre's (see Forsyth, *Diff. Eqns.*, 4th Ed., §§ 202, 203).

The following example illustrates a change of variables that sometimes causes difficulty through failure to notice the precise meaning of the symbols.

Ex. 3. If z is a function $\varphi(x, y, t)$ of three independent variables x, y, t, and if *two* of the variables x, y are changed to two other independent variables u, v by the substitution $x=f(u, v, t)$, $y=g(u, v, t)$, find $\partial z/\partial t$ when z is expressed in terms of u, v, t.

If $\varphi(x, y, t)$ becomes $\psi(u, v, t)$ when the substitution has been made, it is plain that, while $z=\varphi=\psi$, the variable t occurs in $\psi(u, v, t)$ in quite a different way from that in which it appears in $\varphi(x, y, t)$; hence though $\varphi=\psi$ it is not possible (in general) to have $\partial\varphi/\partial t=\partial\psi/\partial t$.

The simplest way of treating this and similar cases is to change *all* the variables; the substitution will then be, if $t=s$,

$$x=f(u, v, s),\ y=g(u, v, s),\ t=s. \qquad (1)$$

Now denote z by φ or by ψ according as the differentiation refers to the old variables x, y, t or to the new u, v, s; by the usual rule we now have

$$\frac{\partial\psi}{\partial s}=\frac{\partial\varphi}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial\varphi}{\partial y}\frac{\partial y}{\partial s}+\frac{\partial\varphi}{\partial t}\frac{\partial t}{\partial s}.$$

From (1) we get $\partial x/\partial s=f_s$, $\partial y/\partial s=g_s$, $\partial t/\partial s=1$ and therefore

$$\frac{\partial\psi}{\partial s}=\frac{\partial\varphi}{\partial x}f_s+\frac{\partial\varphi}{\partial y}g_s+\frac{\partial\varphi}{\partial t}. \qquad (2)$$

We might by finding $\partial\psi/\partial u$, $\partial\psi/\partial v$ complete the transformation in the usual way, but the difficulty to be noticed is that $\partial z/\partial t$, when z is expressed in terms of u, v, t, is *equal to the expression on the right of equation* (2) *after t has been substituted in it for s.* Thus if $x=f(u, v, t)$, $y=g(u, v, t)$ the required value of $\partial z/\partial t$ is

$$\frac{\partial\varphi}{\partial x}\frac{\partial f}{\partial t}+\frac{\partial\varphi}{\partial y}\frac{\partial g}{\partial t}+\frac{\partial\varphi}{\partial t};$$

and the expression is often, indeed usually, written as

$$\frac{\partial z}{\partial x}\frac{\partial f}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial g}{\partial t}+\frac{\partial z}{\partial t},$$

on the understanding that z is $\varphi(x, y, t)$.

If the required value of $\partial z/\partial t$ be denoted by $(\partial z/\partial t)$ we have

$$\left(\frac{\partial z}{\partial t}\right)=\frac{\partial z}{\partial t}+\frac{\partial z}{\partial x}\frac{\partial f}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial g}{\partial t}.$$

In this simple case if the new value of $\partial z/\partial t$ alone is required we may at once apply the rule for differentiating a function of a function; we thus find

$$\left(\frac{\partial z}{\partial t}\right)=\frac{\partial z}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial t}+\frac{\partial z}{\partial t}=\frac{\partial z}{\partial x}\frac{\partial f}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial g}{\partial t}+\frac{\partial z}{\partial t}.$$

50. Elimination of Functions. By the elimination of constants it is possible to form Ordinary Differential Equations (*E.T.* Chap. XX); we shall show by means of examples that

Partial Differential Equations may be derived by the elimination of *functions*. The general theory is quite beyond our limits, and for it reference must be made to works on Differential Equations. The notation p, q, r, s, t for the partial derivatives of z with respect to x, y will often be used in the text and in the Exercises (§ 46, Ex. 2).

Ex. 1. The equation $\dfrac{x-a}{z-c}=f\left(\dfrac{y-b}{z-c}\right)$, where f denotes *any* function (an *arbitrary* function), represents a cone which has the point (a, b, c) as vertex and whose generators are the lines given by $(x-a)/(z-c)=f(t)$, $(y-b)/(z-c)=t$. Show that the differential equation of the cone is

$$z-c=(x-a)p+(y-b)q.$$

Differentiate the given equation with respect to x and y respectively; thus, if $df(t)/dt=f'(t)$,

$$(z-c)-(x-a)p=-(y-b)pf'(t)$$
$$-(x-a)q=\{(z-c)-(y-b)q\}f'(t).$$

Eliminate $f'(t)$ between these equations and we find, after a slight reduction, the equation stated.

It will be noticed that the differential equation is independent of the particular function denoted by the symbol f; the function may be a polynomial or any other function so long as it is differentiable.

Ex. 2. If $z=f(x+ay)$ where a is constant and f is an arbitrary function show that $q=ap$.

Let $x+ay=t$; then $p=f'(t)$ and $q=af'(t)$ so that $q=ap$.

In this case we may prove the converse, namely: if $q=ap$ show that z is an arbitrary function of $x+ay$. Change the variables by the substitution $u=x+ay$, $v=x-ay$ *so that* u *and* v *are independent.*

We find $$p=\frac{\partial z}{\partial u}+\frac{\partial z}{\partial v}, \quad q=a\frac{\partial z}{\partial u}-a\frac{\partial z}{\partial v},$$

so that $q=ap$ becomes $2a\dfrac{\partial z}{\partial v}=0$. Hence z is independent of v and is therefore any function of u or $x+ay$.

Ex. 3. If $z=x\varphi\left(\dfrac{y}{x}\right)+\psi\left(\dfrac{y}{x}\right)$ show that

$$x^2r+2xy\,s+y^2t=0, \quad \text{.............................(i)}$$

and by changing the variables from x, y to u, v where $x=u, y=uv$, prove that equation (i) becomes $\partial^2z/\partial u^2=0$.

Let $y/x=v$ and denote by an accent derivatives with respect to v; then differentiating with respect to x, y we find

$$p=\varphi(v)-\frac{y}{x}\varphi'(v)-\frac{y}{x^2}\psi'(v), \quad q=\varphi'(v)+\frac{1}{x}\psi'(v),$$

so that $$p+\frac{y}{x}q=\varphi(v), \quad \text{.................................(ii)}$$

Again, differentiate (ii) with respect to x, y; thus

$$r+\frac{y}{x}s-\frac{y}{x^2}q=\frac{-y}{x^2}\varphi'(v), \quad s+\frac{y}{x}t+\frac{1}{x}q=\frac{1}{x}\varphi'(v),$$

so that, by elimination of $\varphi'(v)$,

$$x^2r+2xy\,s+y^2t=0. \qquad \text{.............................(i)}$$

Next we have $u=x$, $v=y/x$ and therefore

$$p=\frac{\partial z}{\partial u}-\frac{v}{u}\frac{\partial z}{\partial v}, \quad q=\frac{1}{u}\frac{\partial z}{\partial v},$$

$$r=\frac{\partial p}{\partial u}-\frac{v}{u}\frac{\partial p}{\partial v}=\frac{\partial^2 z}{\partial u^2}-2\frac{v}{u}\frac{\partial^2 z}{\partial u\,\partial v}+\frac{v^2}{u^2}\frac{\partial^2 z}{\partial v^2}+\frac{2v}{u^2}\frac{\partial z}{\partial v},$$

$$s=\frac{\partial q}{\partial x}=\frac{\partial q}{\partial u}-\frac{v}{u}\frac{\partial q}{\partial v}=\frac{1}{u}\frac{\partial^2 z}{\partial u\,\partial v}-\frac{v}{u^2}\frac{\partial^2 z}{\partial v^2}-\frac{1}{u^2}\frac{\partial z}{\partial v},$$

$$t=\frac{\partial q}{\partial y}=\frac{1}{u}\frac{\partial q}{\partial v}=\frac{1}{u^2}\frac{\partial^2 z}{\partial v^2}.$$

Hence equation (i) becomes $u^2\dfrac{\partial^2 z}{\partial u^2}=0$ or $\dfrac{\partial^2 z}{\partial u^2}=0$, the integral of which is $z=Au+B$ *where A, B are constants with respect to u but may be any functions of v*; that is $z=u\varphi(v)+\psi(v)$. Thus equation (i) has $z=x\varphi(y/x)+\psi(y/x)$ as an integral where φ and ψ are arbitrary functions.

Ex. 4. Show that each of the functions defined by the equations

$$z=ax+by+ab, \text{ ....(i) and } z=2\sqrt{(xy)}+1, \text{ ....(ii)}$$

where a and b are any constants, are integrals of the differential equation

$$z=xp+yq+pq. \text{ ....(iii)}$$

From (i), $p=a$, $q=b$, and elimination of a and b gives (iii).

From (ii), $p=\sqrt{(y/x)}$, $q=\sqrt{(x/y)}$ and therefore

$$xp+yq+pq=\sqrt{(xy)}+\sqrt{(xy)}+1=z, \text{ by (ii).}$$

The integral given by (ii) cannot be obtained by assigning particular values to the constants a and b. (Compare *E.T.* p. 432, Ex. 2.)

Note. When it is said that a function is *an integral* of a partial differential equation all that is here meant is that the differential equation may be obtained from the integral by the elimination of *functions*, as in Ex. 3, or of *constants*, as in Ex. 4, (i), or else that, as in Ex. 4, (ii), the differential equation is satisfied by the derivatives of the function, in combination with the expression for the function in terms of the independent variables. Whether, in any given case, more integrals than one exist and, if so, what is the character of the integrals is a subject discussed in works on Differential Equations, *e.g.* Forsyth's *Differential Equations*, Chapters IX and X.

EXERCISES V

1. If $2axz + 2byz + cz^2 = k$, $ax + by + cz = R$, prove that

$$R^3 \frac{\partial^2 z}{\partial x^2} = a^2 k, \quad R^3 \frac{\partial^2 z}{\partial x \partial y} = abk, \quad R^3 \frac{\partial^2 z}{\partial y^2} = b^2 k.$$

2. If $z^3 + 3(ax + by)z = c^3$, prove that

$$x^2 \frac{\partial^2 z}{\partial x^2} + 2xy \frac{\partial^2 z}{\partial x \partial y} + y^2 \frac{\partial^2 z}{\partial y^2} = \frac{2z(ax+by)^2}{(ax+by+z^2)^3}.$$

3. If $ax^3 + by^3 + cz^3 + 3hxyz = k$, show that

$$(hxy + cz^2)^3 \frac{\partial^2 z}{\partial x \partial y} = hk(hxy - cz^2) - 2(abc + h^3)x^2y^2z.$$

4. If $u = \log[\{x + (x^2 - y^2)^{\frac{1}{2}}\}/\{x - (x^2 - y^2)^{\frac{1}{2}}\}]$, then
$du = 2(y\,dx - x\,dy)/y(x^2 - y^2)^{\frac{1}{2}}$.

5. If $u = \cos^{-1}[(1 - xy)/(1 + x^2 + y^2 + x^2y^2)^{\frac{1}{2}}]$, then
$du = dx/(1 + x^2) + dy/(1 + y^2)$.

6. If $xu^2 + 2yv - 4xy = 10$, $yv^2 - 2xu + y^2 = 15$, prove that

$$\frac{\partial u}{\partial x} = \frac{4yv - 2u - u^2 v}{2x(uv+1)}, \quad \frac{\partial v}{\partial x} = \frac{4y + u^2}{2y(uv+1)},$$

$$\frac{\partial u}{\partial y} = \frac{2y + 4xv - v^2}{2x(uv+1)}, \quad \frac{\partial v}{\partial y} = \frac{4x - 2yu - 2v - uv^2}{2y(uv+1)}.$$

7. If $z = a \tan^{-1}(y/x)$ show that

(i) $(1 + q^2)r - 2pqs + (1 + p^2)t = 0$;

(ii) $(rt - s^2)/(1 + p^2 + q^2)^2 = -a^2/(x^2 + y^2 + a^2)^2$.

If $z = a \cosh^{-1}[(x^2 + y^2)^{\frac{1}{2}}/a]$ show that equation (i) holds but that in (ii) $(x^2 + y^2)^2$ must be put in place of $(x^2 + y^2 + a^2)^2$.

8. If $\varphi(x, y, z) = 0$ and $\psi(x, y, z) = 0$ show that, when these equations determine y and z as functions of x,

$$\frac{dy}{dx} = \frac{\partial(\varphi, \psi)}{\partial(z, x)} \Big/ \frac{\partial(\varphi, \psi)}{\partial(y, z)}, \quad \frac{dz}{dx} = \frac{\partial(\varphi, \psi)}{\partial(x, y)} \Big/ \frac{\partial(\varphi, \psi)}{\partial(y, z)}.$$

9. If $\xi = y^2 + z^{-2}$, $\eta = z^2 + x^{-2}$, $\zeta = x^2 + y^{-2}$ and if u is a function of x, y, z prove that

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} + 2\left(\xi\frac{\partial u}{\partial \xi} + \eta\frac{\partial u}{\partial \eta} + \zeta\frac{\partial u}{\partial \zeta}\right)$$
$$= 4\left(y^2\frac{\partial u}{\partial \xi} + z^2\frac{\partial u}{\partial \eta} + x^2\frac{\partial u}{\partial \zeta}\right).$$

10. If $x = a_1\xi + b_1\eta$, $y = a_2\xi + b_2\eta$, $(a_1b_2 \neq a_2b_1)$, and if $f(x, y)$ become $F(\xi, \eta)$ when expressed in terms of ξ, η prove that

$$x_1\frac{\partial f}{\partial x} + y_1\frac{\partial f}{\partial y} = \xi_1\frac{\partial F}{\partial \xi} + \eta_1\frac{\partial F}{\partial \eta},$$

where ξ_1, η_1 are the values of ξ, η when x_1, y_1 are the values of x, y.

Extend the result to n variables $x_1, x_2, \ldots, x_n$ and $\xi_1, \xi_2, \ldots, \xi_n$.

11. If $f(x, y)$ is a polynomial of degree n in x and y and if

$$t^n f(x/t, y/t) = \varphi(x, y, t)$$

prove that
$$x\varphi_x + y\varphi_y + t\varphi_t = n\varphi,$$
and that, if $f(x, y) = 0$,

$$[x\varphi_x + y\varphi_y + t\varphi_t]_{t=1} = 0, \text{ or } xf_x + yf_y + [\varphi_t]_{t=1} = 0.$$

Deduce that the equation of the tangent at (x, y) on the curve $f(x, y) = 0$ is, X, Y being current coordinates,

$$Xf_x + Yf_y + [\varphi_t]_{t=1} = 0.$$

Find also the corresponding form of the equation of the tangent plane at (x, y, z) on the surface $f(x, y, z) = 0$.

12. If $f(x, y) = u_n + u_{n-1} + \ldots + u_1 + u_0$, where u_r is a polynomial in x, y that is homogeneous and of degree r ($u_0 = const.$), show that the equation of the tangent at (x, y) on the curve $f(x, y) = 0$ is

$$Xf_x + Yf_y + u_{n-1} + 2u_{n-2} + \ldots + (n-1)u_1 + nu_0 = 0.$$

If $u_1 = ax + by$, prove that the polar of the origin is

$$aX + bY + nu_0 = 0.$$

Find the corresponding equation of the tangent plane at (x, y, z) on the surface $f(x, y, z) = 0$, where u_r is now a homogeneous polynomial of degree r in x, y, z.

13. Change the variables x, y, z in the equation

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = nu$$

to ξ, η, ζ where $\xi = x/z$, $\eta = y/z$, $\zeta = z$ and show that the equation becomes $\zeta\, \partial u/\partial\zeta = nu$. Deduce that u is of the form $\zeta^n F(\xi, \eta)$, *i.e.* $z^n F(x/z, y/z)$, and is therefore homogeneous.

14. If u is a function of the differences $y - z$, $z - x$, $x - y$ of the independent variables x, y, z, prove that

$$\partial u/\partial x + \partial u/\partial y + \partial u/\partial z = 0. \quad \text{.........................(i)}$$

Deduce by a suitable change of variables a form of equation (i) which shows that u is a function of the differences of x, y, z.

15. If $z = f[(ny - mz)/(nx - lz)]$ where f is an arbitrary function, prove that
$$(nx - lz)p + (ny - mz)q = 0.$$

16. If $z - ax = f(z - by)$, show that

$$bp + aq = ab,$$

and give a geometrical interpretation.

17. If $u = \log r + \theta$ where $r^2 = x^2 + y^2$, $\tan\theta = y/x$ and $z = r\varphi(u)$, the function φ being arbitrary, show that

$$(x + y)p - (x - y)q = z.$$

18. If $u = f(v)$ where u, v are functions of x, y, z, prove that

$$\frac{\partial(u, v)}{\partial(y, z)}\frac{\partial z}{\partial x} + \frac{\partial(u, v)}{\partial(z, x)}\frac{\partial z}{\partial y} = \frac{\partial(u, v)}{\partial(x, y)}.$$

19. If $u=\varphi(x+at)+\psi(x-at)$, φ, ψ arbitrary, show that

$$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}. \qquad \text{(i)}$$

Prove also that if the variables x, t are changed to y, z, where $y=x+at$, $z=x-at$ equation (i) becomes $\partial^2 u/\partial y\,\partial z=0$, and deduce that $u=f(y)+F(z)$, (f, F arbitrary).

20. If z is defined by the equation $y=x\varphi(z)+\psi(z)$, where φ and ψ are arbitrary functions, show that

$$\text{(i) } p+q\varphi(z)=0\,; \quad \text{(ii) } q^2 r-2pqs+p^2 t=0.$$

21. In the differential equation $r-2s+t=0$ change the independent variables to u and v, where $x=u$, $x+y=v$, and prove that the equation becomes $\partial^2 z/\partial u^2=0$.

Deduce that $z=x\varphi(x+y)+\psi(x+y)$ is an integral of the given equation.

22. If $z=f\{x+\varphi(y)\}$, where f and φ are arbitrary functions, prove that

$$ps=qr.$$

23. If z is defined by the equations, φ arbitrary,

$$z\frac{d\varphi(\alpha)}{d\alpha}=\{y-\varphi(\alpha)\}^2, \quad (x+\alpha)\frac{d\varphi(\alpha)}{d\alpha}=y-\varphi(\alpha),$$

prove that $\qquad pq=z.$

24. If $\{z-\varphi(\alpha)\}^2=x^2(y^2-\alpha^2)$ and $\{z-\varphi(\alpha)\}\dfrac{d\varphi(\alpha)}{d\alpha}=\alpha x^2$, show that $pq=xy$.

25. If $z=ax+by+c$ where a, b, c are functions of a variable λ that satisfy the equation

$$x\frac{da}{d\lambda}+y\frac{db}{d\lambda}+\frac{dc}{d\lambda}=0,$$

prove that $rt-s^2=0$. Give a geometrical interpretation.

26. If z is a function of x, y and $x=u+v$, $y=uv$, prove

$$\text{(i) } (u-v)\frac{\partial z}{\partial x}=u\frac{\partial z}{\partial u}-v\frac{\partial z}{\partial v}, \quad (u-v)\frac{\partial z}{\partial y}=\frac{\partial z}{\partial v}-\frac{\partial z}{\partial u}\,;$$

$$\text{(ii)} \qquad \frac{\partial(x, y)}{\partial(u, v)}\cdot\frac{\partial(u, v)}{\partial(x, y)}=1.$$

27. If u and v are functions of x, y defined by the equations

$$v=\frac{\partial F(x, u)}{\partial u}, \quad y=-\frac{\partial F(x, u)}{\partial x},$$

where F is an arbitrary function, prove that

$$\frac{\partial(u, v)}{\partial(x, y)}=1.$$

28. If f and F are two functions of x, y such that

$$\frac{\partial f(x, y)}{\partial x}=\frac{\partial F(x, y)}{\partial y}, \quad \frac{\partial f(x, y)}{\partial y}=-\frac{\partial F(x, y)}{\partial x},$$

prove that if $x=r\cos\theta$, $y=r\sin\theta$,

$$\frac{\partial f}{\partial r}=\frac{1}{r}\frac{\partial F}{\partial \theta}, \quad \frac{1}{r}\frac{\partial f}{\partial \theta}=-\frac{\partial F}{\partial r}.$$

29. Change the variables x and y in the equation

$$y^2r-x^2t=xp-yq$$

to u and v where $u=x^2-y^2$, $v=2xy$, and show that the new equation is

$$\left(u\frac{\partial}{\partial v}-v\frac{\partial}{\partial u}\right)\frac{\partial z}{\partial v}=0.$$

30. If $x=r\cos\theta$, $y=r\sin\theta$, prove that the equation

$$xy\left(\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}\right)-(x^2-y^2)\frac{\partial^2 u}{\partial x\,\partial y}=0$$

becomes

$$r\frac{\partial^2 u}{\partial r\,\partial\theta}-\frac{\partial u}{\partial\theta}=0,$$

and show that $u=r\varphi(\theta)+\psi(r)$, where φ and ψ are arbitrary functions.

31. If X, Y denote respectively the operators

$$X\equiv x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}, \quad Y\equiv x^2\frac{\partial}{\partial x}+y^2\frac{\partial}{\partial y},$$

evaluate

$$X^2(x^my^n), \quad Y^2(x^my^n).$$

If $p=\partial z/\partial x$, $q=\partial z/\partial y$, show that

$$X(x^2p+y^2q)=Y(xp+yq+z).$$

Prove that, if r is any positive integer,

$$YX^r=(X-1)^rY.$$

32. Prove that if in the equation

$$\frac{\partial^2 z}{\partial x^2}+2xy^2\frac{\partial z}{\partial x}+2(y-y^3)\frac{\partial z}{\partial y}+x^2y^2z=0,$$

the variables x, y are changed to u, v, where $x=uv$, $y=1/v$, the new equation is obtained by writing u for x and v for y—that is, z is the same function of u, v as of x, y.

33. If the variables x, y in the equation

$$(x^2+y^2)(r+t)+4xy\,s+2xp+2yq=0$$

are changed to u, v, where $2x=e^u+e^v$, $2y=e^u-e^v$, show that the new equation is

$$\frac{\partial^2 z}{\partial u^2}+\frac{\partial^2 z}{\partial v^2}=0.$$

34. The coordinates of a point P with respect to two sets of rectangular axes with a common origin are (x, y, z) and (ξ, η, ζ), and the direction

cosines of the second set of axes with respect to the first are l_1, m_1, n_1; l_2, m_2, n_2; l_3, m_3, n_3 respectively. If u is a function of x, y, z, prove that

$$\frac{\partial u}{\partial x}=l_1\frac{\partial u}{\partial \xi}+l_2\frac{\partial u}{\partial \eta}+l_3\frac{\partial u}{\partial \zeta},$$

with similar expressions for $\partial u/\partial y$ and $\partial u/\partial z$ so that the operators $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$ and $\partial/\partial \xi$, $\partial/\partial \eta$, $\partial/\partial \zeta$ are changed by the same formulae as the variables x, y, z and ξ, η, ζ.

Deduce:

$$\text{(i)}\ \sum_{x,y,z}\left(\frac{\partial u}{\partial x}\right)^2=\sum_{\xi,\eta,\zeta}\left(\frac{\partial u}{\partial \xi}\right)^2;\quad \text{(ii)}\ \sum_{x,y,z}\frac{\partial^2 u}{\partial x^2}=\sum_{\xi,\eta,\zeta}\frac{\partial^2 u}{\partial \xi^2}$$

where
$$\sum_{x,y,z}\left(\frac{\partial u}{\partial x}\right)^2=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2, \text{ etc.}$$

35. If $r^2=x^2+y^2+z^2$, and if the function $f(r, z)$ satisfies the equation

$$\nabla^2 u\equiv\sum_{x,y,z}\frac{\partial^2 u}{\partial x^2}\equiv\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0,$$

deduce from Example 34 that the function $f(r, lx+my+nz)$ also satisfies the equation, provided $l^2+m^2+n^2=1$.

36. If $x=r\sin\theta\cos\varphi$, $y=r\sin\theta\sin\varphi$, $z=r\cos\theta$, prove that

$$\frac{\partial r}{\partial x}=\sin\theta\cos\varphi,\quad \frac{\partial r}{\partial y}=\sin\theta\sin\varphi,\quad \frac{\partial r}{\partial z}=\cos\theta;$$

$$\frac{\partial \theta}{\partial x}=\frac{\cos\theta\cos\varphi}{r},\quad \frac{\partial \theta}{\partial y}=\frac{\cos\theta\sin\varphi}{r},\quad \frac{\partial \theta}{\partial z}=-\frac{\sin\theta}{r};$$

$$\frac{\partial \varphi}{\partial x}=-\frac{\sin\varphi}{r\sin\theta},\quad \frac{\partial \varphi}{\partial y}=\frac{\cos\varphi}{r\sin\theta},\quad \frac{}{\partial z}=0.$$

Find also the derivatives of x, y, z with respect to r, θ, φ.

37. If u is a function of the independent variables x, y, z, prove, using the values in Example 36, that

$$\frac{\partial u}{\partial x}=\sin\theta\cos\varphi\frac{\partial u}{\partial r}+\frac{\cos\theta\cos\varphi}{r}\frac{\partial u}{\partial \theta}-\frac{\sin\varphi}{r\sin\theta}\frac{\partial u}{\partial \varphi};$$

$$\frac{\partial u}{\partial y}=\sin\theta\sin\varphi\frac{\partial u}{\partial r}+\frac{\cos\theta\sin\varphi}{r}\frac{\partial u}{\partial \theta}+\frac{\cos\varphi}{r\sin\theta}\frac{\partial u}{\partial \varphi};$$

$$\frac{\partial u}{\partial z}=\cos\theta\frac{\partial u}{\partial r}-\frac{\sin\theta}{r}\frac{\partial u}{\partial \theta};$$

$$\sum_{x,y,z}\left(\frac{\partial u}{\partial x}\right)^2=\left(\frac{\partial u}{\partial r}\right)^2+\left(\frac{1}{r}\frac{\partial u}{\partial \theta}\right)^2+\left(\frac{1}{r\sin\theta}\frac{\partial u}{\partial \varphi}\right)^2.$$

38. Using the abbreviations P, Q, R, where

$$P=\sin^2\theta\frac{\partial^2 u}{\partial r^2}+\frac{\cos^2\theta}{r^2}\frac{\partial^2 u}{\partial \theta^2}+\frac{\sin 2\theta}{r}\frac{\partial^2 u}{\partial r\,\partial \theta}+\frac{\cos^2\theta}{r}\frac{\partial u}{\partial r}-\frac{\sin 2\theta}{r^2}\frac{\partial u}{\partial \theta},$$

$$Q=\frac{1}{r}\frac{\partial^2 u}{\partial r\,\partial \varphi}+\frac{\cot\theta}{r^2}\frac{\partial^2 u}{\partial \theta\,\partial \varphi}-\frac{1}{r^2\sin^2\theta}\frac{\partial u}{\partial \varphi},$$

$$R=\frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial \varphi^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{\cot\theta}{r^2}\frac{\partial u}{\partial \theta},$$

prove that

$$\frac{\partial^2 u}{\partial x^2}=P\cos^2\varphi-Q\sin 2\varphi+R\sin^2\varphi,$$

$$\frac{\partial^2 u}{\partial y^2}=P\sin^2\varphi+Q\sin 2\varphi+R\cos^2\varphi,$$

$$\frac{\partial^2 u}{\partial z^2}=\cos^2\theta\,\frac{\partial^2 u}{\partial r^2}+\sin^2\theta\left(\frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2}+\frac{1}{r}\frac{\partial u}{\partial r}\right)-\sin 2\theta\left(\frac{1}{r}\frac{\partial^2 u}{\partial r\,\partial\theta}-\frac{1}{r^2}\frac{\partial u}{\partial\theta}\right)$$

and deduce that $\nabla^2 u$ (*E.T.* p. 238) is equal to

$$\frac{\partial^2 u}{\partial r^2}+\frac{2}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2}+\frac{\cot\theta}{r^2}\frac{\partial u}{\partial\theta}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\varphi^2}\quad\text{...............(i)}$$

$$=\frac{1}{r^3}\left\{r^2\frac{\partial^2(ru)}{\partial r^2}+\frac{\partial^2(ru)}{\partial\theta^2}+\cot\theta\,\frac{\partial(ru)}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2(ru)}{\partial\varphi^2}\right\}.\quad\text{.........(ii)}$$

39. In Example 38 let $ru=k^2v$ and $r\varrho=k^2=const.$; if the variable r is changed to ϱ show that the expression (ii) in Example 38 for $\nabla^2 u$ becomes

$$\frac{\varrho^2}{k^4}\left\{\varrho^2\frac{\partial^2(\varrho v)}{\partial\varrho^2}+\frac{\partial^2(\varrho v)}{\partial\theta^2}+\cot\theta\,\frac{\partial(\varrho v)}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2(\varrho v)}{\partial\varphi^2}\right\}.$$

Hence show that if $F(x, y, z)$ satisfies the equation $\nabla^2 u=0$ so does

$$\frac{1}{r}F\left(\frac{k^2x}{r^2},\ \frac{k^2y}{r^2},\ \frac{k^2z}{r^2}\right).\qquad\text{(Kelvin.)}$$

[If $\xi=\varrho\sin\theta\cos\varphi$, $\eta=\varrho\sin\theta\sin\varphi$, $\zeta=\varrho\cos\theta$ the equation $\nabla^2 u=0$ becomes $\nabla^2 v=0$ where in $\nabla^2 v$ the variables are ξ, η, ζ.

But

$$x=k^2\xi/\varrho^2,\ y=k^2\eta/\varrho^2,\ z=k^2\zeta/\varrho^2,\ \varrho^2=\xi^2+\eta^2+\zeta^2\ \text{and}$$

$$v=\frac{r}{k^2}F(x, y, z)=\frac{1}{\varrho}F\left(\frac{k^2\xi}{\varrho^2},\ \frac{k^2\eta}{\varrho^2},\ \frac{k^2\zeta}{\varrho^2}\right).$$

In v we may now put x, y, z, r in place of $\xi, \eta, \zeta, \varrho$.]

40. If $\xi=k^2x/r^2$, $\eta=k^2y/r^2$, $\zeta=k^2z/r^2$, $r^2=x^2+y^2+z^2$, $k^2=$const., and if Σ is a summation as to x, y, z, show that:

(i) $\Sigma\eta_x\zeta_x\equiv\eta_x\zeta_x+\eta_y\zeta_y+\eta_z\zeta_z=0,\ \Sigma\zeta_x\xi_x=0,\ \Sigma\xi_x\eta_x=0$;

(ii) $\Sigma(\xi_x)^2=\Sigma(\eta_x)^2=\Sigma(\zeta_x)^2=k^4/r^4$;

(iii) $\Sigma\xi_{xx}=-2\xi/r^2,\ \Sigma\eta_{xx}=-2\eta/r^2,\ \Sigma\zeta_{xx}=-2\zeta/r^2$.

If $f(x, y, z)=\varphi(\xi, \eta, \zeta)$;

(iv) $xf_x+yf_y+zf_z=-(\xi\varphi_\xi+\eta\varphi_\eta+\zeta\varphi_\zeta)$.

41. If $z=F(x, y)$ is the equation of a surface and if $x=f(u, v)$, $y=g(u, v)$, $z=h(u, v)$ (See Bell, *Coord. Geom. of Three Dimensions*, § 185), prove that

$$\frac{\partial z}{\partial x}=-\frac{J_1}{J_3},\quad\frac{\partial z}{\partial y}=-\frac{J_2}{J_3},$$

where

$$J_1=\frac{\partial(y, z)}{\partial(u, v)},\quad J_2=\frac{\partial(z, x)}{\partial(u, v)},\quad J_3=\frac{\partial(x, y)}{\partial(u, v)}.$$

42. $\varphi(x, y, z;\ p, q, r)$ is a function of two sets of independent variables x, y, z and p, q, r, three in each set, and φ is homogeneous and of the second degree in p, q, r. The variables p, q, r are changed to ξ, η, ζ by the transformation

$$\xi = \frac{\partial \varphi}{\partial p}, \quad \eta = \frac{\partial \varphi}{\partial q}, \quad \zeta = \frac{\partial \varphi}{\partial r}; \qquad \text{.(i)}$$

if $\varphi(x, y, z;\ p, q, r)$ becomes $\psi(x, y, z;\ \xi, \eta, \zeta)$ prove that

$$\left.\begin{aligned} p = \frac{\partial \psi}{\partial \xi}, \quad q = \frac{\partial \psi}{\partial \eta}, \quad r = \frac{\partial \psi}{\partial \zeta} \\ \frac{\partial \varphi}{\partial x} = -\frac{\partial \psi}{\partial x}, \quad \frac{\partial \varphi}{\partial y} = -\frac{\partial \psi}{\partial y}, \quad \frac{\partial \varphi}{\partial z} = -\frac{\partial \psi}{\partial z} \end{aligned}\right\} \qquad \text{.........(ii)}$$

[By Euler's Theorem, $p\dfrac{\partial \varphi}{\partial p} + q\dfrac{\partial \varphi}{\partial q} + r\dfrac{\partial \varphi}{\partial r} = 2\varphi$ so that we may write

$$p\xi + q\eta + r\zeta = 2\varphi = \varphi + \psi.$$

Now take the complete differential of each side and apply (i); the results follow at once.

If there are two sets of n independent variables, $x_1, x_2, \ldots, x_n$ and $p_1, p_2, \ldots, p_n$, and if $\varphi(x_1, \ldots, x_n;\ p_1 \ldots, p_n)$ is homogeneous and of the second degree in $p_1, \ldots, p_n$ the transformation

$$\xi_r = \frac{\partial \varphi}{\partial p_r},\ r = 1, 2, \ldots, n;\ \varphi(x_1, \ldots, x_n;\ p_1, \ldots, p_n) \equiv \psi(x_1, \ldots, x_n;\ \xi_1, \ldots, \xi_n)$$

gives
$$p_r = \frac{\partial \psi}{\partial \xi_r}, \quad \frac{\partial \varphi}{\partial x_r} = -\frac{\partial \psi}{\partial x_r}\Big].$$

CHAPTER V

IMPLICIT FUNCTIONS. JACOBIANS

51. Implicit Functions. Throughout our work it has been assumed that an equation $f(x, y)=0$ determines y as a function of x; y may be determined for all real values of x, or only for a limited range of x, and the equation may define more than one value of y. For example, the equation

$$10x^2 - 6xy + y^2 - 1 = 0$$

gives $\qquad y = 3x + \sqrt{(1-x^2)}$ and $y = 3x - \sqrt{(1-x^2)}$,

and thus defines two functions, each of which exists for the range $-1 \leqq x \leqq 1$.

It is seldom, however, that it is possible to obtain, as in this simple case, the expression of y as an explicit function of x, and appeal is made to the representation of the equation by a curve as sufficient evidence for the existence of a function y that can be treated as if it were explicitly defined as a function of x. Exercise in the tracing of curves from their equations is from this point of view of special value as it produces what may be called a practical certainty that, at least in a very large number of cases, the equation does define y as a single-valued, or as a many-valued, function of x. It is desirable, however, that the advanced student should investigate the question more closely, and we now consider *Existence Theorems*—that is, theorems that specify conditions which guarantee that an equation does define a function, even though the actual determination of an analytical expression for the function may demand new processes or may be, from a practical standpoint, too laborious. For many purposes, however, it is the fact that an equation does define a function, rather than an expression for the function thus defined, that is of real importance; hence the value of Existence Theorems.

In the following discussion the conditions imposed on the functions are more drastic than is absolutely necessary but they allow sufficient scope for the demands of ordinary analysis; for an excellent treatment of the whole subject, with less drastic conditions, the student is referred to Goursat's *Cours d'Analyse*, Vol. I, Chap. III.

In regard to the general character of the discussion it may be helpful to the student to note that the existence of a function is only established, in the first place, for a small range of each of the variables. Thus if $f(x, y, z)$ is zero for the values a, b, c of x, y, z respectively it is shown that, under certain conditions, the equation $f(x, y, z)=0$ determines z as a function $\varphi(x, y)$ of x and y where x, y, z differ but little (in general) from a, b, c. In the language of geometry the point (x, y, z) is confined to a region (R_1) defined by such inequalities as

$$| x-a | \leqq h, \ | y-b | \leqq k, \ | z-c | \leqq l$$

where h, k, l are positive and may be very "small." It may be possible afterwards to extend the range of the variables for which $\varphi(x, y)$ exists, but the essential element of the discussion is the proof of the existence of $\varphi(x, y)$ for values of x, y that differ but little from a, b respectively.

52. Existence Theorem I. *Let $f(x, y)$ be a function of the two variables x, y which satisfies the conditions:* (i) *$f(x, y)$ is zero for $x=a$, $y=b$;* (ii) *the partial derivatives f_x and f_y exist and are continuous near (a, b);* (iii) *the derivative f_y is not zero for $x=a$, $y=b$. When these conditions are fulfilled there is one and only one function y of x, say $y=\varphi(x)$, that satisfies the equation $f(x, y)=0$ for every x that is near a,* and that is equal to b when $x=a$; further $\varphi(x)$ has a derivative $\varphi'(x)$ and both $\varphi(x)$ and $\varphi'(x)$ are continuous.*

It may be noted first that $f(x, y)$ is continuous near (a, b) since the derivatives f_x and f_y exist.

Let the neighbourhood of (a, b) for which conditions (ii) and (iii) hold be the region (R_1) defined by the inequalities

$$|x-a| \leqq h_1, \ \ |y-b| \leqq k_1 \ldots\ldots\ldots\ldots\ldots\ldots (R_1)$$

* This means that the equation $f\{x, \phi(x)\}=0$ is an *identity* if x is near a. A similar meaning is to be given to satisfying an equation in the corresponding statements in the other theorems.

If $|h| \leqq h_1$ and $|k| \leqq k_1$ we have by the Mean Value Theorem, § 45, since $f(a, b)=0$,

$$f(a+h, b+k)=hf_x(a+\theta h, b+\theta k)+kf_y(a+\theta h, b+\theta k) \quad \ldots\ldots(1)$$

where $0<\theta<1$.

Now f_y is continuous near (a, b) and is not zero for $x=a, y=b$; there is therefore a neighbourhood of (a, b) in which f_y is not zero, and it will be assumed that h_1 and k_1 have been chosen so that f_y is not zero in (R_1). In (R_1) the continuous function f_y has therefore always the same sign as it has at (a, b), and is numerically greater than some constant B, say $|f_y| > B$. Again f_x is continuous and therefore $|f_x|$ is bounded in (R_1), say $|f_x|<A$, a constant.

Next let h_2 be the smaller of the two numbers h_1 and Bk_1/A, so that $Ah_2 \leqq Bk_1$, and let (R_2) be the region defined by the inequalities

$$|x-a| \leqq h_2, \; |y-b| \leqq k_1 \quad \ldots\ldots\ldots\ldots\ldots\ldots(R_2)$$

In equation (1) suppose $|h| \leqq h_2$ and $k=\pm k_1$; the second term on the right of (1) will then be numerically greater than Bk_1 while the first term will be numerically less than Ah_2 so that the sign of the right side of (1), and therefore the sign of $f(a+h, b+k)$, will be that of the second term on the right of (1). Now f_y has always the same sign and therefore if the second term is positive when $k=k_1$ it is negative when $k=-k_1$, while if it is negative when $k=k_1$ it is positive when $k=-k_1$. Hence if h, or x where $x=a+h$, is kept constant the function $f(x, y)$ changes sign as y varies from $b-k_1$ to $b+k_1$ and therefore, since $f(x, y)$ is a continuous function of y, it is zero for at least one value of y in the interval $(b-k_1, b+k_1)$. Further, $f(x, y)$ is not zero for more than one value of y in that interval; for, if it were, f_y would by Rolle's Theorem vanish at least once in the interval and it does not. Thus for every value of x in the interval $(a-h_2, a+h_2)$ there is one and only one value of y that satisfies the equation $f(x, y)=0$; since to each value of x there corresponds one and only one value of y we may denote this value of y by $\varphi(x)$ and obviously y or $\varphi(x)$ is equal to b when x is equal to a, since a is an admissible value of x.

Again, if $y=\varphi(x)$ and $y+k=\varphi(x+h)$ both $f(x, y)$ and $f(x+h, y+k)$ are zero provided (x, y) and $(x+h, y+k)$ are

both in (R_2), and equation (1) holds if for a we put x and for b we put y; hence, since $f(x+h, y+k)$ is zero, we have

$$\frac{\varphi(x+h)-\varphi(x)}{h}=\frac{k}{h}=-\frac{f_x(x+\theta h,\ y+\theta k)}{f_y(x+\theta h,\ y+\theta k)}$$

so that
$$\varphi'(x)=\mathop{\mathcal{L}}_{h\to 0} k/h=-f_x(x,\ y)/f_y(x,\ y).$$

Since f_x and f_y are continuous, and f_y is not zero, $\varphi'(x)$ exists and is continuous; $\varphi(x)$ is therefore also continuous.

It is proved, therefore, that when the conditions of the theorem are fulfilled the equation $f(x, y)=0$ determines y as a function $\varphi(x)$ of x that exists at least for the range $|x-a|\leqq h_2$.

The range of x for which $\varphi(x)$ has been proved to exist may, however, be greater than that just found. Suppose that (a', b') is a point in (R_2) for which $y=\varphi(x)$; near (a', b') the derivatives f_x and f_y are continuous, f_y is not zero and $f(a', b')=0$ so that the conditions of the theorem hold for a certain neighbourhood of (a', b'), say for a region (R_3) defined by the inequalities

$$|x-a'|\leqq h_3,\ |y-b'|\leqq k_3 \quad\ldots\ldots\ldots\ldots\ldots\ldots(R_3)$$

It may happen, and usually does, that the region (R_3) projects beyond the region (R_2) so as to contain points for which x is greater than $a+h_2$ (or less than $a-h_2$). Now in the region (R_3) the equation $f(x, y)=0$ determines one and only one value of y, say $y=\psi(x)$, for which $b'=\psi(a')$ and $f(x, y)=0$; in the part common to (R_2) and (R_3), $y=\varphi(x)$ and $b'=\varphi(a')$ so that in the common part $\psi(x)$ is the same function as $\varphi(x)$. We may, therefore, in the part of (R_3) that projects beyond (R_2), denote $\psi(x)$ by $\varphi(x)$; in this way we see that the range for which the unique solution of $f(x, y)=0$ exists may be extended. By taking a suitable point (a'', b'') in (R_3) it may be possible similarly to extend the range still further; this procedure will only be stopped when it is impossible to find in the last region reached a point, (α, β) say, that provides a new region in which all the conditions of the theorem are fulfilled.

The process thus briefly described is called *Analytical Continuation*. The Existence Theorem is quite independent of the process of continuation; this process is mentioned

simply to show that the range for which $\varphi(x)$ exists may be much wider than is given by the proof which has been developed above. A similar process of continuation is applicable in regard to the other Existence Theorems but will not be further referred to.

53. Derivatives of Implicit Function. When it is known that the equation $f(x, y)=0$ defines y as a function of x that has a derivative dy/dx that derivative may of course be obtained simply by differentiating the equation with respect to x, on the understanding that y is a function $\varphi(x)$ of x. Thus we find

$$f_x + f_y \frac{dy}{dx} = 0 \quad \text{.........................(1)}$$

If the higher partial derivatives of $f(x, y)$ are continuous we obtain the higher derivatives of y or $\varphi(x)$ by successive differentiations of (1), provided always that f_y is not zero; thus

$$f_{xx} + f_{xy}\frac{dy}{dx} + \left(f_{xy} + f_{yy}\frac{dy}{dx}\right)\frac{dy}{dx} + f_y \frac{d^2y}{dx^2} = 0,$$

or

$$f_{xx} + 2f_{xy}\frac{dy}{dx} + f_{yy}\left(\frac{dy}{dx}\right)^2 + f_y \frac{d^2y}{dx^2} = 0. \quad \text{.........(2)}$$

Provided f_y is not zero this equation determines the second derivative, and in a similar way the third and higher derivatives may be found.

54. Existence Theorem II. *Let $f(x_1, x_2, \dots, x_n, y)$ be a function of $n+1$ variables $x_1, x_2, \dots, x_n, y$ which satisfies the conditions:* (i) *$f(x_1, x_2, \dots, x_n, y)$ is zero for $x_1=a_1, x_2=a_2, \dots, x_n=a_n, y=b$;* (ii) *all the partial derivatives $f_{x_1}, f_{x_2}, \dots, f_{x_n}, f_y$ exist and are continuous near $(a_1, a_2, \dots, a_n, b)$;* (iii) *the derivative f_y is not zero for $x_1=a_1, x_2=a_2, \dots, x_n=a_n, y=b$. When these conditions are fulfilled there is one and only one function y of $x_1, x_2, \dots, x_n$, say $y=\varphi(x_1, x_2, \dots, x_n)$, that satisfies the equation $f(x_1, x_2, \dots, x_n, y)=0$ for every point $(x_1, x_2, \dots, x_n)$ near $(a_1, a_2, \dots, a_n)$ and that is equal to b when $x_1=a_1, \dots, x_n=a_n$; further, the derivatives $\varphi_{x_1}, \varphi_{x_2}, \dots, \varphi_{x_n}$ exist and are continuous so that φ is also continuous.*

The proof is essentially a repetition of that given for Theorem I, and may therefore be developed with less detail.

Let the neighbourhood of $(a_1, a_2, \ldots, a_n, b)$ for which conditions (ii) and (iii) hold be the region (R_1) defined by the inequalities

$$|x_1 - a_1| \leqq h_1, |x_2 - a_2| \leqq h_1, \ldots, |x_n - a_n| \leqq h_1, |y - b| \leqq k_1 \ldots (R_1)$$

where h_1 and k_1 are positive. As in the proof of Theorem I it may be assumed that h_1 and k_1 have been chosen so that in (R_1) the derivative f_y is numerically greater than B and each of the derivatives $f_{x_1}, f_{x_2}, \ldots, f_{x_n}$ numerically less than A where A and B are positive constants. It is also to be noted that f_y does not change sign in (R_1).

Next, in view of the application of the Mean Value Theorem, let h_2 be the smaller of the two numbers h_1 and Bk_1/nA so that $nAh_2 \leqq Bk_1$ and let (R_2) be the region defined by the inequalities

$$|x_1 - a_1| \leqq h_2, |x_2 - a_2| \leqq h_2, \ldots, |x_n - a_n| \leqq h_2, |y - b| \leqq k_1 \ldots (R_2)$$

Suppose now that $|\delta_1| \leqq h_2, |\delta_2| \leqq h_2, \ldots, |\delta_n| \leqq h_2, |k| \leqq k_1$; we have by the Mean Value Theorem, since $f(a_1, a_2, \ldots, a_n, b) = 0$,

$$\begin{aligned} &f(a_1 + \delta_1, a_2 + \delta_2, \ldots, a_n + \delta_n, b + k) \\ &= \delta_1 \bar{f}_{x_1} + \delta_2 \bar{f}_{x_2} + \ldots + \delta_n \bar{f}_{x_n} + k\bar{f}_y \quad \ldots\ldots\ldots\ldots\ldots(1) \end{aligned}$$

where the derivatives are all taken for the values

$$a_1 + \theta\delta_1, a_2 + \theta\delta_2, \ldots, a_n + \theta\delta_n, b + \theta k \quad (0 < \theta < 1).$$

This equation (1) is treated as in the corresponding equation in the proof of Theorem I. Let $k = \pm k_1$; then the last term on the right of (1) is numerically greater than the numerical value of the sum of the first n terms on the right of (1). For $k_1 |\bar{f}_y|$ is greater than $k_1 B$, while

$$|\delta_1 \bar{f}_{x_1} + \delta_2 \bar{f}_{x_2} + \ldots + \delta_n \bar{f}_{x_n}| < h_2 (A + A + \ldots + A),$$

that is, $< nAh_2$, so that, by the value of h_2, the term $k\bar{f}_y$ is numerically greater than the numerical value of the sum of the first n terms.

Hence when $k = \pm k_1$ the sign of the right side of (1) and therefore the sign of $f(a_1 + \delta_1, \ldots, a_n + \delta_n, b + k)$ is that of the last term on the right of (1). As before, if

$$x_r = a_r + \delta_r, \quad r = 1, 2, \ldots, n,$$

and if x_r is kept constant, we see that $f(x_1, x_2, \ldots, x_n, y)$ changes sign once and only once as y varies from $b - k_1$ to $b + k_1$, and

therefore the equation $f(x_1, x_2, \dots, x_n, y)=0$ determines one and only one value of y as a function of $x_1, x_2, \dots, x_n$, say $y=\varphi(x_1, x_2, \dots, x_n)$; further $y=b$ when

$$x_1=a_1,\ x_2=a_2, \dots, x_n=a_n$$

since $a_1, a_2, \dots, a_n$ are admissible values of x.

Again, if

$$y=\varphi(x_1, x_2, \dots, x_n),\ y+k=\varphi(x_1+\delta_1, x_2+\delta_2, \dots, x_n+\delta_n)$$

where $(x_1, x_2 \dots, x_n, y)$, $(x_1+\delta_1, x_2+\delta_2, \dots, x_n+\delta_n, y+k)$ lie in (R_2), the Mean Value Theorem gives the equation

$$0=\delta_1\bar{f}_{x_1}+\delta_2\bar{f}_{x_2}+\dots+\delta_n\bar{f}_{x_n}+k\bar{f}_y$$

where the derivatives are taken for the values

$$x_1+\theta\delta_1, \dots, x_n+\theta\delta_n, y+\theta k \quad (0<\theta<1).$$

Hence if each δ, except δ_r, is zero we find

$$\frac{\varphi(x_1, \dots, x_r+\delta_r, \dots, x_n)-\varphi(x_1, \dots, x_r \dots, x_n)}{\delta_r}=\frac{k}{\delta_r}=-\frac{\bar{f}_{x_r}}{\bar{f}_y}$$

so that $$\partial\varphi/\partial x_r=-f_{x_r}/f_y.$$

Since f_y is not zero and the partial derivatives of f are continuous, the function φ has continuous derivatives with respect to $x_1, x_2, \dots, x_n$; φ is itself continuous since each of its partial derivatives is continuous.

The higher derivatives of φ may be obtained as before by differentiating the equation $f=0$ when the derivatives of f satisfy the usual conditions; it should be specially noted that f_y must not be zero.

The range of the variables $x_1, x_2, \dots, x_n$ for which the function φ has been shown to exist may usually be extended by the process of analytical continuation sketched in connection with Theorem I.

In the proof of Theorem III the determinant called the Jacobian plays an important part; we therefore define it and prove one or two of its properties before taking up Theorem III.

55. The Jacobian. Let $y_r=f_r(x_1, x_2, \dots, x_n)$, $r=1, 2, \dots, n$, be n functions of the variables $x_1, x_2, \dots, x_n$, and let each of

the partial derivatives $\partial y_r/\partial x_s$ be a continuous function of $x_1, x_2, \ldots, x_n$. The determinant J

where
$$J = \begin{vmatrix} \dfrac{\partial y_1}{\partial x_1}, & \dfrac{\partial y_1}{\partial x_2}, & \ldots, & \dfrac{\partial y_1}{\partial x_n} \\ \dfrac{\partial y_2}{\partial x_1}, & \dfrac{\partial y_2}{\partial x_2}, & \ldots, & \dfrac{\partial y_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial y_n}{\partial x_1}, & \dfrac{\partial y_n}{\partial x_2}, & \ldots, & \dfrac{\partial y_n}{\partial x_n} \end{vmatrix}$$

is called the *Jacobian* or the *Functional Determinant* of the functions $y_1, y_2, \ldots, y_n$ with respect to $x_1, x_2, \ldots, x_n$ and is denoted by

$$\frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} \quad \text{or} \quad \frac{d(y_1, y_2, \ldots, y_n)}{d(x_1, x_2, \ldots, x_n)}$$

or
$$J\begin{pmatrix} y_1, y_2, \ldots, y_n \\ x_1, x_2, \ldots, x_n \end{pmatrix} \quad \text{or} \quad J(y_1, y_2, \ldots, y_n).$$

For $n=1$ the determinant is simply dy_1/dx_1, the derivative of y_1 with respect to x_1; the first of the notations given for J is suggested by a certain analogy between the properties of the Jacobian and the derivative, as shown by the following theorem.

If $z_1, z_2, \ldots, z_n$ are functions of $y_1, y_2, \ldots, y_n$ and $y_1, y_2, \ldots, y_n$ are functions of $x_1, x_2, \ldots, x_n$ then

$$\frac{\partial(z_1, z_2, \ldots, z_n)}{\partial(x_1, x_2, \ldots, x_n)} = \frac{\partial(z_1, z_2, \ldots, z_n)}{\partial(y_1, y_2, \ldots, y_n)} \cdot \frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, y_n)} \quad \ldots\ldots(1)$$

If $n=1$ this is the usual relation $\dfrac{dz_1}{dx_1} = \dfrac{dz_1}{dy_1}\dfrac{dy_1}{dx_1}$. The proof is simply a theorem in the multiplication of determinants combined with the rule for the derivative of a function of a function.

Find the product of the two determinants on the right of (1) by the "row by column" rule; that is, to find the element in the rth row and the sth column of the product, multiply the elements in the rth row of the first determinant by the corresponding elements in the sth column of the second and add the products. We have

$$r\text{th row} \quad \frac{\partial z_r}{\partial y_1}, \frac{\partial z_r}{\partial y_2}, \ldots, \frac{\partial z_r}{\partial y_n},$$

$$s\text{th column} \quad \frac{\partial y_1}{\partial x_s}, \frac{\partial y_2}{\partial x_s}, \ldots, \frac{\partial y_n}{\partial x_s},$$

so that the element in the rth row and the sth column of the product is

$$\frac{\partial z_r}{\partial y_1}\frac{\partial y_1}{\partial x_s}+\frac{\partial z_r}{\partial y_2}\frac{\partial y_2}{\partial x_s}+\dots+\frac{\partial z_r}{\partial y_n}\frac{\partial y_n}{\partial x_s},$$

and this is equal to $\partial z_r/\partial x_s$ which is the element in the rth row and the sth column of the Jacobian of $z_1, z_2, \dots, z_n$.

If we suppose $z_r = x_r$, $r=1, 2, \dots, n$, and assume that the equations which define $y_1, y_2, \dots, y_n$ as functions of $x_1, x_2, \dots, x_n$ determine, conversely, $x_1, x_2, \dots, x_n$ as functions of $y_1, y_2, \dots, y_n$, we find

$$\frac{\partial(y_1, y_2, \dots, y_n)}{\partial(x_1, x_2, \dots, x_n)}\cdot\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(y_1, y_2, \dots, y_n)}=1 \quad\dots\dots\dots\dots(2)$$

because $\dfrac{\partial(x_1, x_2, \dots, x_n)}{\partial(x_1, x_2, \dots, x_n)}=1$, since $\partial x_r/\partial x_s=0$ unless $r=s$, in which case it is equal to 1.

The theorem expressed in (1) is a particular form of the following: *If $y_1, y_2, \dots, y_n$ are determined as functions of $x_1, x_2, \dots, x_n$ by the equations $\varphi_r(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n)=0$, $r=1, 2, \dots, n$, then*

$$\frac{\partial(\varphi_1, \varphi_2, \dots, \varphi_n)}{\partial(x_1, x_2, \dots, x_n)}=(-1)^n\frac{\partial(\varphi_1, \varphi_2, \dots, \varphi_n)}{\partial(y_1, y_2, \dots, y_n)}\cdot\frac{\partial(y_1, y_2, \dots, y_n)}{\partial(x_1, x_2, \dots, x_n)}. \quad (3)$$

Differentiation of the equation $\varphi_r=0$ with respect to x_s gives

$$\frac{\partial\varphi_r}{\partial x_s}+\frac{\partial\varphi_r}{\partial y_1}\frac{\partial y_1}{\partial x_s}+\frac{\partial\varphi_r}{\partial y_2}\frac{\partial y_2}{\partial x_s}+\dots+\frac{\partial\varphi_r}{\partial y_n}\frac{\partial y_n}{\partial x_s}=0,$$

so that the element in the rth row and the sth column of the determinant which is the product of the two determinants on the right of (3) is $-\partial\varphi_r/\partial x_s$ from which the result follows.

Again, if $y_{m+1}, y_{m+2} \dots, y_n$ are constant with respect to $x_1, x_2, \dots, x_m$, or if $y_1, y_2, \dots, y_m$ are constant with respect to $x_{m+1}, x_{m+2}, \dots, x_n$, then

$$\frac{\partial(y_1, \dots y_m, y_{m+1}, \dots, y_n)}{\partial(x_1, \dots x_m, x_{m+1}, \dots, x_n)}=\frac{\partial(y_1, \dots, y_m)}{\partial(x_1, \dots, x_m)}\cdot\frac{\partial(y_{m+1} \dots, y_n)}{\partial(x_{m+1} \dots, x_n)}\dots(4)$$

and in particular

$$\frac{\partial(y_1 \dots y_m, x_{m+1} \dots, x_n)}{\partial(x_1 \dots x_m, x_{m+1} \dots, x_n)}=\frac{\partial(y_1 \dots, y_m)}{\partial(x_1 \dots, x_m)} \quad\dots\dots\dots\dots(5)$$

To prove, note that $\partial y_r/\partial x_s = 0$ if y_r is constant with respect to x_s.

Ex. If ξ, η are functions of the three variables x, y, z and x, y, z are functions of the two independent variables u, v, prove that

$$\frac{\partial(\xi, \eta)}{\partial(u, v)} = \frac{\partial(\xi, \eta)}{\partial(x, y)}\frac{\partial(x, y)}{\partial(u, v)} + \frac{\partial(\xi, \eta)}{\partial(y, z)}\frac{\partial(y, z)}{\partial(u, v)} + \frac{\partial(\xi, \eta)}{\partial(z, x)}\frac{\partial(z, x)}{\partial(u, v)}.$$

56. Existence Theorem III. *Let $f_r(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n)$, where r takes the values $1, 2, \dots, n$, be n functions of the $m+n$ variables $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ which satisfy the following conditions:* (i) *each of the n functions $f_1, f_2, \dots, f_n$ is zero for*

$$x_1 = a_1, \; x_2 = a_2, \qquad x_m = a_m, \; y_1 = b_1, \; y_2 = b_2, \dots, y_n = b_n,$$

that is, at the point $(a_1, \dots, a_m, b_1, \dots, b_n)$; (ii) *all the first partial derivatives of the functions $f_1, f_2, \dots, f_n$ with respect to*

$$x_1, \qquad y_1, \quad , y_n$$

exist and are continuous near $(a_1, \dots, a_m, b_1, \dots, b_n)$; (iii) *the Jacobian J of the functions $f_1, f_2, \dots, f_n$ with respect to $y_1, y_2, \dots, y_n$ is not zero at $(a_1, \dots, a_m, b_1, \dots, b_n)$.*

When these conditions are fulfilled there is one and only one system of functions $y_1, y_2, \dots, y_n$ of the variables $x_1, x_2, \dots, x_m$, say

$$y_1 = \varphi_1(x_1, x_2, \dots, x_m), \; y_2 = \varphi_2(x_1, x_2, \dots, x_m), \dots \\ y_n = \varphi_n(x_1, x_2, \dots, x_m),$$

such that, for all points $(x_1, x_2, \dots, x_m)$ near $(a_1, a_2, \dots, a_m)$, they satisfy the n equations $f_1 = 0, f_2 = 0, \dots, f_n = 0$ and become equal to $b_1, b_2, \dots, b_n$ respectively when $x_1, x_2, \dots, x_m$ are equal respectively to $a_1, a_2, \dots, a_m$; further, all the first partial derivatives of $\varphi_1, \varphi_2, \dots, \varphi_n$ with respect to $x_1, x_2, \dots, x_m$ exist and are continuous so that $\varphi_1, \varphi_2, \dots, \varphi_n$ are also continuous

If $n = 1$, that is, if there is only one function, Theorem III is simply Theorem II, so that it is true when there is only one function. It will be proved to hold for n functions by showing that if it holds for $(n-1)$ functions, say the functions $f_2, f_3, \dots, f_n$, it will hold for the n functions $f_1, f_2, \dots, f_n$. It will thus be assumed that the equations $f_2 = 0, f_3 = 0, \dots, f_n = 0$ determine $y_2, y_3 \dots, y_n$ in terms of $x_1, x_2, \dots x_m$ and y_1, these functions of $x_1, \dots x_m, y_1$ satisfying all the conditions respecting initial

values, derivatives and continuity stated in Theorem III; the proof for n functions will then be developed.

The Jacobian J is

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial y_1}, & \dfrac{\partial f_1}{\partial y_2}, & \dfrac{\partial f_1}{\partial y_3}, & \dots, & \dfrac{\partial f_1}{\partial y_n} \\ \dfrac{\partial f_2}{\partial y_1}, & \dfrac{\partial f_2}{\partial y_2}, & \dfrac{\partial f_2}{\partial y_3}, & \dots, & \dfrac{\partial f_2}{\partial y_n} \\ & & & & \\ \dfrac{\partial f_n}{\partial y_1}, & \dfrac{\partial f_n}{\partial y_2}, & \dfrac{\partial f_n}{\partial y_3}, & \dots, & \dfrac{\partial f_n}{\partial y_n} \end{vmatrix}$$

Since J is not zero at $(a_1, \dots, a_m, b_1 \dots, b_n)$ the co-factors of the elements in the first column cannot all be zero; it will be assumed that the notation has been chosen so that the co-factor, J_1 say, of $\partial f_1/\partial y_1$ is not zero. J_1 is the Jacobian of the functions $f_2, f_3, \dots, f_n$ with respect to $y_2, y_3, \dots, y_n$ *and is not zero* at $(a_1, \dots, a_m, b_1, \dots, b_n)$.

The conditions of Theorem III are fulfilled by the functions $f_2, f_3, \dots, f_n$ and determine $y_2, y_3, \dots, y_n$ as functions of $x_1, x_2, \dots, x_m, y_1$ because the hypothesis is that the theorem holds for $(n-1)$ functions. (It is to be noted that at this stage y_1 is associated with $x_1, x_2, \dots, x_m$ and the solutions for $y_2, \dots, y_n$ involve y_1 as well as $x_1, \dots, x_m$). Hence we have

$$y_r = \psi_r(x_1, x_2, \dots, x_m, y_1), \quad r=2, 3, \dots n, \quad \text{............(1)}$$

where $\psi_2, \psi_3, \dots, \psi_n$ satisfy $f_2=0, f_3=0, \dots, f_n=0$, take the values $b_2, b_3, \dots, b_n$ respectively at $(a_1, \dots, a_m, b_1)$ and are, as well as their first partial derivatives, continuous near $(a_1, \dots, a_m, b_1)$.

Let ψ_r be substituted for y_r in f_1 and let

$$f_1(x_1, \dots, x_m, y_1, \psi_2, \dots, \psi_n) = F_1(x_1, \dots, x_m, y_1). \quad \text{......(2)}$$

It will now be shown that $F_1=0$ determines y_1 as a function of $x_1, \dots, x_m$. Of the conditions required by Theorem II the first is fulfilled since $F_1(a_1, \dots, a_m, b_1)$ is equal to

$$f_1(a_1, \dots, a_m, b_1, b_2, \dots, b_n)$$

and is therefore zero. Again the partial derivatives of F_1 satisfy the second condition; for, by the rule for differentiating a function of a function, we have, since

$F_1 = f_1(x_1, \dots, x_m, y_1, y_2, \dots, y_n)$ where $y_2, y_3, \dots, y_n$ are the functions $\psi_2, \psi_3, \dots, \psi_n$

$$\frac{\partial F_1}{\partial y_1} = \frac{\partial f_1}{\partial y_1} + \frac{\partial f_1}{\partial y_2}\frac{\partial \psi_2}{\partial y_1} + \frac{\partial f_1}{\partial y_3}\frac{\partial \psi_3}{\partial y_1} + \dots + \frac{\partial f_1}{\partial y_n}\frac{\partial \psi_n}{\partial y_1} \quad \dots\dots\dots\dots\dots(3)$$

$$\frac{\partial F_1}{\partial x_s} = \frac{\partial f_1}{\partial x_s} + \frac{\partial f_1}{\partial y_2}\frac{\partial \psi_2}{\partial x_s} + \frac{\partial f_1}{\partial y_3}\frac{\partial \psi_3}{\partial x_s} + \dots + \frac{\partial f_1}{\partial y_n}\frac{\partial \psi_n}{\partial x_s}, \quad s = 1, 2, \dots m,$$

and all the derivatives on the right are continuous.

Lastly, the third condition that $\partial F_1/\partial y_1$ is not zero at $(a_1, \dots, a_m, b_1)$ is also fulfilled, as will now be shown.

When the values of y_r given by equation (1) are substituted in $f_2, f_3, \dots, f_n$ these functions vanish *identically* (that is, for all values of $x_1, x_2, \dots, x_m, y_1$) near $(a_1, \dots, a_m, b_1)$ and therefore their derivatives with respect to y_1 are zero. Hence

$$0 = \frac{\partial f_2}{\partial y_1} + \frac{\partial f_2}{\partial y_2}\frac{\partial \psi_2}{\partial y_1} + \frac{\partial f_2}{\partial y_3}\frac{\partial \psi_3}{\partial y_1} + \dots + \frac{\partial f_2}{\partial y_n}\frac{\partial \psi_n}{\partial y_1}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad \dots\dots\dots\dots(4)$$

$$0 = \frac{\partial f_n}{\partial y_1} + \frac{\partial f_n}{\partial y_2}\frac{\partial \psi_2}{\partial y_1} + \frac{\partial f_n}{\partial y_3}\frac{\partial \psi_3}{\partial y_1} + \dots + \frac{\partial f_n}{\partial y_n}\frac{\partial \psi_n}{\partial y_1}.$$

Now multiply the 2nd column of J by $\partial \psi_2/\partial y_1$, the 3rd by $\partial \psi_3/\partial y_1, \dots$, the nth by $\partial \psi_n/\partial y_1$ and add to the first column; this transformation makes no change in the value of J. The first element in the first column of J as thus transformed is $\partial F_1/\partial y_1$, by the first of equations (3), while all the other elements of the first column are zero, as shown by equations (4); hence the transformed determinant is equal to the product of $\partial F_1/\partial y_1$ and the co-factor J_1 of $\partial f_1/\partial y_1$ in J so that

$$J = J_1 \frac{\partial F_1}{\partial y_1},$$

and as J and J_1 are both different from zero at $(a_1, \dots, a_m, b_1, \dots, b_n)$ so is $\partial F_1/\partial y_1$. Condition (iii) of Theorem II is therefore satisfied.

Thus the equation $F_1 = 0$ gives

$$y_1 = \varphi_1(x_1, x_2, \dots, x_m),$$

and if now in equation (1) this value of y_1 is substituted we find

$$y_r = \psi_r(x_1, \dots, x_m, \varphi_1) = \varphi_r(x_1, \dots, x_m), \quad r = 2, 3, \dots, n.$$

At $(a_1, \dots, a_m)$ the functions $\varphi_1, \varphi_2, \dots, \varphi_n$ are equal to $b_1, b_2, \dots, b_n$ respectively, while near $(a_1, \dots, a_m)$ the first partial derivatives of the functions are continuous.

It has therefore been proved that if Theorem III is true for $(n-1)$ functions it is true for n functions; since the theorem is true for one function it is therefore true in general.

Ex. **Inversion.** Let the n functions $f_1, f_2, \ldots, f_n$ of the n independent variables $x_1, x_2, \ldots, x_n$ and also all the first partial derivatives of the functions be continuous; prove that if the Jacobian J of $f_1, f_2, \ldots, f_n$ with respect to $x_1, x_2, \ldots, x_n$ is not identically zero the n equations

$$y_r = f_r(x_1, x_2, \ldots, x_n), \quad r = 1, 2, \ldots n \quad \text{.......................(1)}$$

determine, inversely, $x_1, x_2, \ldots, x_n$ as functions of $y_1, y_2, \ldots, y_n$.

Let $F_r(x_1, x_2, \ldots, x_n, y_r) = f_r(x_1, x_2, \ldots, x_n) - y_r$ and we have a case of Theorem III. The Jacobian of $F_1, \ldots, F_n$ with respect to $x_1, \ldots, x_n$ is the same as that of $f_1, f_2, \ldots, f_n$ with respect to $x_1, \ldots, x_n$ (the *notation* differs from that of Theorem III by the interchange of x and y).

By hypothesis the Jacobian of $F_1, \ldots, F_n$ with respect to $x_1, \ldots, x_n$, which is independent of $y_1, \ldots, y_n$, is not identically zero, and therefore, there is a set of values $x_1 = a_1, x_2 = a_2, \ldots, x_n = a_n$ for which it is not zero. For these values of $x_1, \ldots, x_n$ let $y_r = f_r(a_1, a_2, \ldots, a_n) = b_r$, $r = 1, 2, \ldots, n$. The functions F_r satisfy the conditions of Theorem III. For, (i) $F_r = 0$ at $(a_1, \ldots, a_n, b_1, \ldots, b_n)$; (ii) $\partial F_r/\partial x_s = \partial f_r/\partial x_s$ and is therefore continuous while $\partial F_r/\partial y_s = -1$ if $r = s$ but $= 0$ if $r \neq s$; (iii) J is not zero near $(a_1, \ldots, a_n, b_1, \ldots, b_n)$. Hence the *equations* $F_r = 0$, that is, $y_r = f_r(x_1, \ldots, x_n)$, give the system

$$x_r = \varphi_r(y_1, \ldots, y_n) \quad r = 1, 2, \ldots, n.$$

This example is the problem of *Inversion*; the functions $x_1, \ldots, x_n$ are inverse to the functions $y_1, \ldots, y_n$.

The values $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are often called the "initial values" of the variables $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ respectively.

57. Dependence of Functions. Let $f_1, f_2, \ldots, f_n$ be n functions of m independent variables $x_1, x_2, \ldots, x_m$. The functions are said to be *dependent* if they satisfy one or more equations in which the variables $x_1, x_2, \ldots, x_m$ do not appear explicitly—equations therefore which are satisfied whatever be the values of $x_1, x_2, \ldots, x_m$; the functions are said to be *independent* if they do not satisfy any equation of the kind just mentioned, that is, an equation in which $x_1, x_2, \ldots, x_m$ do not appear explicitly.

Again, it may be said that the functions are independent if it is impossible to express one of them in terms of the others. For example, if f_1, f_2, f_3 are the functions

$$x^2 + y^2 + z^2, \quad xy + xz + yz, \quad x + y + z$$

respectively they are dependent since $f_1 + 2f_2 = f_3^2$, an equation in which x, y, z do not appear explicitly; here $f_1 = f_3^2 - 2f_2$

so that one function may be expressed in terms of the rest. Of course the functions would still be dependent if one could be expressed in terms of some (not necessarily all) of the others.

In the following discussion we give the method of Goursat, *Cours d'Analyse*, Vol. I, Chap. III, to which the student is referred for fuller information; the treatment by Bateman, *Differential Equations*, Chap. VI is also instructive.

THEOREM. *Let the n functions $f_1, f_2, \dots, f_n$ of the n independent variables $x_1, x_2, \dots, x_n$ and also all their first partial derivatives be continuous; the necessary and sufficient condition that the functions should be dependent is that their Jacobian with respect to $x_1, x_2, \dots, x_n$ should be identically zero.*

(i) The condition is necessary. Let J be the Jacobian and

$$y_r = f_r(x_1, x_2, \dots, x_n), \quad r = 1, 2, \dots, n.$$

If J is not identically zero it is possible (§ 56, *example*) to determine $x_1, x_2, \dots, x_n$ so that

$$x_r = \varphi_r(y_1, y_2, \dots, y_n), \quad r = 1, 2, \dots, n$$

where $y_1, y_2, \dots, y_n$ may have *any* values near their respective initial values $b_1, b_2, \dots, b_n$. Since the values of $y_1, y_2, \dots, y_n$ are quite arbitrary, it is impossible that they can satisfy an equation $F(y_1, y_2, \dots, y_n) = 0$ in which the coefficients are constants, that is, independent of $x_1, x_2, \dots, x_n$. If, therefore, J is not identically zero the functions are independent.

(ii) The condition is sufficient. It will secure brevity and at the same time show quite plainly the lines of the proof for n variables to take $n = 4$; the notation will also be simplified.

Let the independent variables be x, y, z, t, and the functions u, v, w, s where

$$u = f_1(x, y, z, t), v = f_2(x, \dots, t), w = f_3(x, \dots, t), s = f_4(x, \dots, t); \dots(1)$$

the Jacobian J will be

$$J = \begin{vmatrix} \dfrac{\partial f_1}{\partial x}, & \dfrac{\partial f_1}{\partial y}, & \dfrac{\partial f_1}{\partial z}, & \dfrac{\partial f_1}{\partial t} \\[2ex] \dfrac{\partial f_2}{\partial x}, & \dfrac{\partial f_2}{\partial y}, & \dfrac{\partial f_2}{\partial z}, & \dfrac{\partial f_2}{\partial t} \\[2ex] \dfrac{\partial f_3}{\partial x}, & \dfrac{\partial f_3}{\partial y}, & \dfrac{\partial f_3}{\partial z}, & \dfrac{\partial f_3}{\partial t} \\[2ex] \dfrac{\partial f_4}{\partial x}, & \dfrac{\partial f_4}{\partial y}, & \dfrac{\partial f_4}{\partial z}, & \dfrac{\partial f_4}{\partial t} \end{vmatrix}$$

It is now assumed that J is *identically* zero; different cases arise, dependent on the minors of J.

I. Suppose that the first minors of J are not all zero; we may suppose that the notation has been so chosen that the minor, J_1 say, obtained by deleting the 4th row and the 4th column of J is not zero. This minor is the Jacobian of f_1, f_2, f_3 with respect to x, y, z;

$$J_1 = \frac{\partial(f_1, f_2, f_3)}{\partial(x, y, z)}.$$

Now, since J_1 is not zero, the first three of equations (1) may, by Existence Theorem III, be solved for x, y, z in terms of u, v, w, t; let the solutions be

$$x = \varphi_1(u, v, w, t),\ y = \varphi_2(u, v, w, t),\ z = \varphi_3(u, v, w, t). \quad \ldots(2)$$

When these values are substituted in the fourth of equations (1) we get

$$s = f_4(\varphi_1, \varphi_2, \varphi_3, t) = F(u, v, w, t). \quad \ldots\ldots\ldots\ldots\ldots(3)$$

It is to be noted that u, v, w, t are independent variables and that the first three of equations (1) become identities when $\varphi_1, \varphi_2, \varphi_3$ are substituted for x, y, z respectively.

It will now be shown that $\partial F/\partial t$ is zero, so that F does not contain t explicitly. We have, by differentiating $F(u, v, w, t)$ that is, $f_4(x, y, z, t)$ where x, y, z are the functions $\varphi_1, \varphi_2, \varphi_3$ respectively

$$\frac{\partial F}{\partial t} = \frac{\partial f_4}{\partial x}\frac{\partial \varphi_1}{\partial t} + \frac{\partial f_4}{\partial y}\frac{\partial \varphi_2}{\partial t} + \frac{\partial f_4}{\partial z}\frac{\partial \varphi_3}{\partial t} + \frac{\partial f_4}{\partial t}. \quad \ldots\ldots\ldots\ldots\ldots(4)$$

Again, by differentiating the first three of equations (1) which are identities when $\varphi_1, \varphi_2, \varphi_3$ are put in place of x, y, z respectively

$$\begin{aligned} 0 &= \frac{\partial f_1}{\partial x}\frac{\partial \varphi_1}{\partial t} + \frac{\partial f_1}{\partial y}\frac{\partial \varphi_2}{\partial t} + \frac{\partial f_1}{\partial z}\frac{\partial \varphi_3}{\partial t} + \frac{\partial f_1}{\partial t} \\ 0 &= \frac{\partial f_2}{\partial x}\frac{\partial \varphi_1}{\partial t} + \frac{\partial f_2}{\partial y}\frac{\partial \varphi_2}{\partial t} + \frac{\partial f_2}{\partial z}\frac{\partial \varphi_3}{\partial t} + \frac{\partial f_2}{\partial t}. \quad \ldots\ldots\ldots\ldots(5) \\ 0 &= \frac{\partial f_3}{\partial x}\frac{\partial \varphi_1}{\partial t} + \frac{\partial f_3}{\partial y}\frac{\partial \varphi_2}{\partial t} + \frac{\partial f_3}{\partial z}\frac{\partial \varphi_3}{\partial t} + \frac{\partial f_3}{\partial t} \end{aligned}$$

Equations (5) determine the t-derivatives of $\varphi_1, \varphi_2, \varphi_3$; but it is not necessary actually to solve them since the t-derivatives of $\varphi_1, \varphi_2, \varphi_3$ can be eliminated by transforming J, thus:

Multiply the 1st column of J by $\partial\varphi_1/\partial t$, the 2nd by $\partial\varphi_2/\partial t$, the 3rd by $\partial\varphi_3/\partial t$ and add to the 4th ; the elements in the 1st, 2nd and 3rd rows of the 4th column will now be zero, by equations (5), while the element in the 4th row is $\partial F/\partial t$, by equation (4). We thus find, since the value of J is not changed,

$$J = J_1 \frac{\partial F}{\partial t},$$

and therefore $\partial F/\partial t = 0$ since $J=0$, $J_1 \neq 0$, so that F does not contain t explicitly. Hence equation (3) gives $s = F(u, v, w)$, an equation that does not contain x, y, z, t explicitly ; thus the functions u, v, w, s are dependent.

It may be observed that there cannot be a second relation, say $s = F_1(u, v, w)$, that is distinct from $s = F(u, v, w)$; if there were there would be an equation $F = F_1$ connecting u, v, w, and therefore J_1 would be zero, contrary to the hypothesis.

II. Suppose all the first but not all the second minors of J to be zero ; we may assume that a non-zero second minor is that obtained by deleting the 3rd and 4th rows and the 3rd and 4th columns of J. This minor, J_2 say, is the Jacobian of f_1, f_2 with respect to x, y ; since $J_2 \neq 0$, the first two of equations (1) may be solved for x, y, giving, say,

$$x = \psi_1(u, v, z, t), \quad y = \psi_2(u, v, z, t).$$

When these values are substituted in the last two of equations (1) we find

$$w = f_3(\psi_1, \psi_2, z, t) = F_1(u, v, z, t),$$

$$s = f_4(\psi_1, \psi_2, z, t) = F_2(u, v, z, t).$$

It may be shown as before that neither z not t occurs explicitly in F_1 or in F_2. For example,

$$\frac{\partial F_1}{\partial z} = \frac{\partial f_3}{\partial x}\frac{\partial \psi_1}{\partial z} + \frac{\partial f_3}{\partial y}\frac{\partial \psi_2}{\partial z} + \frac{\partial f_3}{\partial z};$$

further, by differentiating the first two of equations (1), which are identities when x, y are replaced by ψ_1, ψ_2 respectively, we find

$$0 = \frac{\partial f_1}{\partial x}\frac{\partial \psi_1}{\partial z} + \frac{\partial f_1}{\partial y}\frac{\partial \psi_2}{\partial z} + \frac{\partial f_1}{\partial z},$$

$$0 = \frac{\partial f_2}{\partial x}\frac{\partial \psi_1}{\partial z} + \frac{\partial f_2}{\partial y}\frac{\partial \psi_2}{\partial z} + \frac{\partial f_2}{\partial z}.$$

By now treating the determinant J_1 on the same lines as was done in the case of J and $\partial F/\partial t$ we see that

$$J_1 = J_2 \frac{\partial F_1}{\partial z}; \ \frac{\partial F_1}{\partial z} = 0 \text{ since } J_1 = 0, \ J_2 \neq 0.$$

To evaluate $\partial F_1/\partial t$ and $\partial F_2/\partial t$ take, not J_1 but, the minor of J which is obtained by deleting the 4th row and the 3rd column of J; this minor is also zero, by hypothesis.

We thus obtain *two* relations $w = F_1(u, v)$, $s = F_2(u, v)$ when the first minors of J are zero but not all the second minors.

III. Suppose all the second (and therefore all the first) minors of J to be zero but not all the third minors. When there are four functions the third minors are simply the elements of J. If we suppose $\partial f_1/\partial x$ not zero we deduce $x = \varphi(u, y, z, t)$ and it is then proved as before that when φ is substituted for x in the other three of equations (1) the variables y, z, t do not occur explicitly so that now there are *three* relations between the functions.

The procedure is clearly general. When there are n functions there is one relation when the first minors of J are not all zero, two relations when the first minors are all zero, but the second minors not all zero, and so on.

If J is zero merely because one or more of the functions $f_1, f_2, \ldots, f_n$ is zero (or constant), it does not follow that the functions are dependent. It must be specially noticed that the proofs assume that J is *identically* zero. The following simple example is usually given.

Let $u = x^2 + y^2 - 1$, $v = x \cos \alpha + y \sin \alpha - 1$, where α is constant. Here $J = 2(x \sin \alpha - y \cos \alpha)$ and is therefore not identically zero. But it is easily seen that

$$u - v^2 - 2v = (\tfrac{1}{2}J)^2,$$

and $J = 0$ if $u = 0$ and $v = 0$; u and v are, however, independent and the relation $u = 0$ is not a consequence of $v = 0$.

Of course, if the Jacobian of $f_1, f_2, \ldots, f_n$ with respect to $x_1, x_2, \ldots, x_n$ is identically zero, and if these functions contain other variables, say $z_1, z_2, \ldots, z_m$, these variables $z_1, \ldots, z_m$ will appear (usually) in the equation or equations that connect $f_1, f_2, \ldots, f_n$; what the theorem just proved guarantees is only that $x_1, \ldots, x_n$ do not appear in these equations.

58. The Hessian. An important Jacobian is that in which the functions $f_1, f_2, \dots, f_n$ are the first partial derivatives $\partial f/\partial x_1, \partial f/\partial x_2, \dots, \partial f/\partial x_n$ of a function $f(x_1, x_2, \dots, x_n)$ of the n independent variables $x_1, x_2, \dots, x_n$; this Jacobian is called the *Hessian* of the function $f(x_1, \dots, x_n)$ and may be denoted by the symbol H_f. The element in the rth row and sth column of the determinant H_f is

$$\frac{\partial}{\partial x_s} \cdot \frac{\partial f}{\partial x_r}, \text{ that is, } \frac{\partial^2 f}{\partial x_s \partial x_r}.$$

Ex. 1. If f is the quadratic form

$$a_{11}x_1^2 + \dots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \dots + 2a_{n-1,n}x_{n-1}x_n,$$

the Hessian of f is a numerical multiple of the discriminant of f; it does not contain any of the variables.

Ex. 2. If $f(x, y, z)$ is a polynomial that is homogeneous and of the nth degree in x, y, z, prove that the Hessian is homogeneous and of degree $3(n-2)$.

EXERCISES VI.

1. If $x = r\cos\theta$, $y = r\sin\theta$, prove $\dfrac{\partial(x, y)}{\partial(r, \theta)} = r$, and if

$$x = r\sin\theta\cos\varphi,\ y = r\sin\theta\sin\varphi,\ z = r\cos\theta,$$

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)} = r^2\sin\theta.$$

2. If $x + y + z = u$, $y + z = uv$, $z = uvw$,

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = u^2 v.$$

If $x_1 + x_2 + \dots + x_n = y_1$, $x_2 + x_3 + \dots + x_n = y_1y_2, \dots$

$$x_r + x_{r+1} + \dots + x_n = y_1y_2 \dots y_r, \dots, x_n = y_1y_2 \dots y_n,$$

$$\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(y_1, y_2, \dots, y_n)} = y_1^{n-1}y_2^{n-2} \dots y_{n-2}^2 y_{n-1}.$$

3. If $y_1 = \cos x_1$, $y_2 = \sin x_1 \cos x_2$, $y_3 = \sin x_1 \sin x_2 \cos x_3$,

$$\frac{\partial(y_1, y_2, y_3)}{\partial(x_1, x_2, x_3)} = (-1)^3 \sin^3 x_1 \sin^2 x_2 \sin x_3,$$

extend the theorem to the case of n functions.

4. If $x = a\cosh\xi\cos\eta$, $y = a\sinh\xi\sin\eta$,

$$\frac{\partial(x, y)}{\partial(\xi, \eta)} = \tfrac{1}{2}a^2(\cosh 2\xi - \cos 2\eta).$$

5. If $x=\sin\theta(1-c^2\sin^2\varphi)^{\frac{1}{2}}$, $y=\cos\theta\cos\varphi$,

$$\frac{\partial(x, y)}{\partial(\theta, \varphi)}=\frac{-\sin\varphi\{(1-c^2)\cos^2\theta+c^2\cos^2\varphi\}}{(1-c^2\sin^2\varphi)^{\frac{1}{2}}}.$$

6. If $u=x/(1-r^2)^{\frac{1}{2}}$, $v=y/(1-r^2)^{\frac{1}{2}}$, $w=z/(1-r^2)^{\frac{1}{2}}$ where $r^2=x^2+y^2+z^2$, show that

$$\frac{\partial(u, v, w)}{\partial(x, y, z)}=\frac{1}{(1-r^2)^{\frac{5}{2}}}.$$

Extend to the case of n functions $y_1, y_2, \dots, y_n$ where

$$y_s=x_s/(1-r^2)^{\frac{1}{2}},\quad r^2=x_1^2+x_2^2+\dots+x_n^2.$$

7. Prove that the functions $3x+2y-z$, $x-2y+z$ and $x(x+2y-z)$ are not independent and find the equation that connects them.

8. The functions u, v, w of x, y, z become functions U, V, W of ξ, η, ζ when x, y, z are changed to ξ, η, ζ by the substitution

$$x=l_1\xi+m_1\eta+n_1\zeta,\quad y=l_2\xi+m_2\eta+n_2\zeta,\quad z=l_3\xi+m_3\eta+n_3\zeta;$$

if M is the determinant $|\, l_1\, m_2\, n_3 \,|$ of the coefficients of the substitution, show that

$$\frac{\partial(U, V, W)}{\partial(\xi, \eta, \zeta)}=M\frac{\partial(u, v, w)}{\partial(x, y, z)}.$$

Extend to the case of n functions of n variables.

9. If $f(x, y, z)$ becomes $F(\xi, \eta, \zeta)$ when the variables are changed as in Example 8, prove that

$$H_F=M^2H_f$$

where H_f and H_F are the Hessians of f and F respectively.

Show that the theorem holds for any number of variables.

10. If $f(x, y, t)$ is homogeneous, of the nth degree, in x, y, t so that $f(x, y, t)=t^nf(x/t, y/t, 1)$, prove that

$$H_f=\begin{vmatrix} f_{xx}, & f_{xy}, & f_{xt} \\ f_{xy}, & f_{yy}, & f_{yt} \\ f_{xt}, & f_{yt}, & f_{tt} \end{vmatrix}=\frac{(n-1)^2}{t^2}\begin{vmatrix} f_{xx}, & f_{xy}, & f_x \\ f_{xy}, & f_{yy}, & f_y \\ f_x, & f_y, & nf/(n-1) \end{vmatrix}$$

If $\xi=x/t$, $\eta=y/t$ and $f(\xi, \eta, 1)=\varphi(\xi, \eta)$, show that

$$H_f=(n-1)^2t^{3(n-2)}\begin{vmatrix} \varphi_{\xi\xi}, & \varphi_{\xi\eta}, & \varphi_\xi \\ \varphi_{\xi\eta}, & \varphi_{\eta\eta}, & \varphi_\eta \\ \varphi_\xi, & \varphi_\eta, & n\varphi/(n-1) \end{vmatrix}$$

11. If u, v are two polynomials in x, y that are homogeneous and of the nth degree, prove that

$$u\,dv-v\,du=\frac{1}{n}\frac{\partial(u, v)}{\partial(x, y)}(x\,dy-y\,dx).$$

12. If $y_r = u_r/u$, $r = 1, 2, \dots, n$, and if u and u_r are functions of the n independent variables $x_1, x_2, \dots, x_n$, prove that

$$\frac{\partial(y_1, y_2, \dots, y_n)}{\partial(x_1, x_2, \dots, x_n)} = \frac{1}{u^{n+1}} \begin{vmatrix} u, & \dfrac{\partial u}{\partial x_1}, & \dfrac{\partial u}{\partial x_2}, & \dots & \dfrac{\partial u}{\partial x_n} \\ u_1, & \dfrac{\partial u_1}{\partial x_1}, & \dfrac{\partial u_1}{\partial x_2}, & \dots & \dfrac{\partial u_1}{\partial x_n} \\ u_2, & \dfrac{\partial u_2}{\partial x_1}, & \dfrac{\partial u_2}{\partial x_2}, & \dots & \dfrac{\partial u_2}{\partial x_n} \\ & & & & \\ u_n, & \dfrac{\partial u_n}{\partial x_1}, & \dfrac{\partial u_n}{\partial x_2}, & \dots & \dfrac{\partial u_n}{\partial x_n} \end{vmatrix}$$

If $u = v/t$ and $u_r = v_r/t$, show that the value of the Jacobian is obtained by substituting v, v_r for u, u_r in the expression on the right and state how the determinants in u and v are connected.

13. If λ, μ, ν are the roots of the equation in k

$$x/(a+k) + y/(b+k) + z/(c+k) = 1,$$

prove that

$$\frac{\partial(x, y, z)}{\partial(\lambda, \mu, \nu)} = -\frac{(\mu-\nu)(\nu-\lambda)(\lambda-\mu)}{(b-c)(c-a)(a-b)}.$$

14. Given that $x = f(u, v, w)$, $y = g(u, v, w)$, $z = h(u, v, w)$ and that J is not zero, where $J = \dfrac{\partial(x, y, z)}{\partial(u, v, w)}$; if, when x, y, z are taken as the coordinates of a point referred to rectangular axes, the three surfaces

$$u = \text{const.}, \quad v = \text{const.}, \quad w = \text{const.},$$

intersect orthogonally, show that $J = \pm \rho_1\rho_2\rho_3$ where

$$\rho_1^2 = f_u^2 + g_u^2 + h_u^2, \quad \rho_2^2 = f_v^2 + g_v^2 + h_v^2, \quad \rho_3^2 = f_w^2 + g_w^2 + h_w^2.$$

[Note that the direction cosines of the normals to the three surfaces are proportional respectively to

$$f_u, g_u, h_u; \quad f_v, g_v, h_v; \quad f_w, g_w, h_w.$$

Also, since $u =$ const. and $v =$ const. intersect orthogonally,

$$f_u f_v + g_u g_v + h_u h_v = 0;$$

similarly $\qquad f_v f_w + \dots + \dots = 0, \; f_w f_u + \dots + \dots = 0.$

The square of the determinant J is $\rho_1^2\rho_2^2\rho_3^2$.]

15. If u, v, w are functions of x, y, z, prove that the rate of variation of u per unit of length along the line of intersection of the surfaces $v = \textit{const.}$, $w = \textit{const.}$ is the quotient of the Jacobian of u, v, w with respect to x, y, z by

$$(v_x^2 + v_y^2 + v_z^2)^{\frac{1}{2}}(w_x^2 + w_y^2 + w_z^2)^{\frac{1}{2}} \sin\theta$$

where θ is the angle at which the surfaces $v = \textit{const.}$ and $w = \textit{const.}$ intersect at (x, y, z).

16. If the Jacobian of n functions of n independent variables is not identically zero, show that the notation of functions and variables may be so chosen that the functions may be represented by

$$y_r = f_r(x_1, x_2, \ldots, x_n), \quad r = 1, 2, \ldots, n,$$

while none of the Jacobians

$$\frac{\partial(f_1, f_2, \ldots, f_r)}{\partial(x_1, x_2, \ldots, x_r)}, \quad r = 1, 2, \ldots, n,$$

is identically zero. Then prove that we may write

$$y_1 = \varphi_1(x_1, x_2, \ldots, x_n), \quad y_2 = \varphi_2(y_1, x_2, x_3, \ldots, x_n),$$
$$y_3 = \varphi_3(y_1, y_2, x_3, \ldots, x_n), \ \ldots, \ y_n = \varphi_n(y_1, y_2, \ldots, y_{n-1}, x_n),$$

and deduce that

$$\frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} = \frac{\partial\varphi_1}{\partial x_1}\frac{\partial\varphi_2}{\partial x_2}\frac{\partial\varphi_3}{\partial x_3}\cdots\frac{\partial\varphi_n}{\partial x_n}.$$

17. Prove that if the functions $f_1(x_1, x_2, \ldots, x_n)$ and $f_2(x_1, x_2, \ldots, x_n)$ are to be connected by an equation in which none of the variables $x_1, x_2, \ldots, x_n$ appears explicitly, it is necessary and sufficient that the corresponding partial derivatives $\partial f_1/\partial x_r$ and $\partial f_2/\partial x_r$, $r = 1, 2, \ldots, n$, should be proportional.

18. The roots of the equation in λ

$$(\lambda - u)^3 + (\lambda - v)^3 + (\lambda - w)^3 = 0$$

are x, y, z; prove that

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = -2\,\frac{(v-w)(w-u)(u-v)}{(y-z)(z-x)(x-y)}$$

CHAPTER VI

INFINITE SERIES. COMPLEX FUNCTIONS OF A REAL VARIABLE

59. Infinite Series. It is necessary, in view of applications in later chapters, to supplement the sketch of Infinite Series given in the *Elementary Treatise* and to discuss briefly the theory of Infinite Products. An excellent treatment both of Series and of Products will be found in Bromwich's Treatise on *Infinite Series*, and the student should consult that book for further developments.

Derangement of Terms. The sum of a finite number of terms is the same in whatever order the terms be taken in calculating the sum, but the word " sum " as applied to the " sum of an infinite series " is not a " sum " in the same sense as that of the " sum of a finite number of terms " ; it is the *limit* of a sum of a finite number of terms and in the case of infinite series the commutative law of addition is not true unless under certain restrictions.

Let Σa_n and Σb_n be two infinite series ; if every term that occurs in one series occurs once and only once in the other, the one series is said to be formed from the other by a *derangement* of the terms of the other series or, simply, to be a derangement of the other series. The following theorem, usually called Dirichlet's Theorem, will now be proved.

DIRICHLET'S THEOREM. *The sum of an absolutely convergent series is the same in whatever order the terms are taken ; the sum of a series that is not absolutely convergent may be changed by a change of the order in which the terms are taken.*

(1) Let the terms be all positive and let Σb_n be a derangement of Σa_n. If $s_n = a_1 + a_2 + \ldots + a_n$, $\sigma_n = b_1 + b_2 + \ldots + b_n$ and if $s_n \to s$ when $n \to \infty$ it has to be proved that $\sigma_n \to s$ when $n \to \infty$.

Since every term of Σb_n occurs in Σa_n it is possible to take n so large that every term in σ_m is a term in s_n and therefore $\sigma_m \leqq s_n$ so that $\sigma_m < s$, a fixed number. Hence σ_m (or σ_n) tends to a limit that cannot exceed s; in other words, Σb_n is convergent and its sum, σ say, is less than or equal to s.

We may now reverse the process. Σb_n is known to be convergent, with a sum σ; Σa_n is a derangement of Σb_n and therefore is convergent and has a sum not less than σ. But the sum of Σa_n is s so that, from the two parts, we have $\sigma \leqq s$ and $s \leqq \sigma$ and therefore $\sigma = s$.

(2) Suppose there is an infinite number both of positive and of negative terms in Σa_n. (If there were only a *finite* number of terms of the one kind these could be neglected so far as the question of convergence is concerned (*E.T.* p. 380, Note) and the series would fall under case (1)). Let P_μ be the sum of the positive terms and $-Q_\nu$ the sum of the negative terms in s_n; then $\mu + \nu = n$ and when n tends to infinity so do μ and ν.

Now Σa_n is absolutely convergent and

$$|a_1| + |a_2| + \ldots + |a_n| = P_\mu + Q_\nu, \quad s_n = P_\mu - Q_\nu,$$

so that both P_μ and Q_ν tend to limits, P and Q say, when μ and ν tend to infinity, and if s is the sum of Σa_n then $s = P - Q$. But the series ΣP_μ and ΣQ_ν are series of positive terms and no derangement of their terms alters their sum. Hence s, the sum of Σa_n, is not altered by any derangement of the terms of Σa_n.

If Σa_n is convergent but not absolutely convergent both of the series ΣP_μ and ΣQ_ν are divergent. For, if $s = \Sigma a_n$, $(P_\mu - Q_\nu)$ tends to s while $(P_\mu + Q_\nu)$ tends to $+\infty$ when $n \to \infty$. If we suppose that, for example, P_μ tends to a limit P when $n \to \infty$ then Q_ν, which is equal to $P_\mu - s_n$ would also tend to a limit, namely $P - s$, and this is impossible since $P_\mu + Q_\nu$ tends to infinity. Hence

$$s = \mathop{\mathcal{L}}_{n\to\infty} \Sigma a_n = \mathop{\mathcal{L}}_{n\to\infty} (P_\mu - Q_\nu),$$

but s is not equal to $\displaystyle \mathop{\mathcal{L}}_{\mu\to\infty} P_\mu - \mathcal{L}\, Q_\nu$;

the difference is $\infty - \infty$, a meaningless expression.

Of course it has not been proved, nor is it the case, that *every* derangement of terms produces an alteration in the sum of a *conditionally* convergent series, as a non-absolutely convergent series is often called, the reason for the name "conditionally convergent" being now obvious. The typical example of a series whose sum may be changed by derangement of its terms is the usual series for log 2; see Exercises II. Exs. 10 and 11. On the general theory see Bromwich, *Infinite Series* (2nd Ed.), pp. 74-77.

60. Tests of Convergence. The following, known as Kummer's Test, is of wide application, the terms of Σa_n being *positive.*

Kummer's Test. *Let Σa_n be a series of positive terms and (d_n) a sequence of positive numbers such that the series $\Sigma(1/d_n)$ is divergent; further, let g_n be defined by the equation*

$$g_n = d_n \frac{a_n}{a_{n+1}} - d_{n+1}.$$

The series Σa_n converges if there is an integer m such that $g_n > \alpha > 0$ when $n \geqq m$, but diverges if there is an integer m such that $g_n < -\alpha < 0$ when $n \geqq m$ (α a positive constant).

Suppose first that $g_n > \alpha > 0$ if $n \geqq m$; then since $a_{n+1} > 0$

$$d_n a_n - d_{n+1} a_{n+1} > \alpha a_{n+1}.$$

In this inequality put $n+1, n+2, \dots, (n+p-1)$ successively in place of n and add corresponding members of the p inequalities; then

$$d_n a_n - d_{n+p} a_{n+p} > \alpha(a_{n+1} + a_{n+2} + \dots + a_{n+p}),\ n \geqq m.$$

The expression on the right side of this inequality is positive; therefore the expression on the left side is also positive and, further, it is less than $d_n a_n$. Hence, putting m in place of n we find that, whatever integer p may be,

$$a_{m+1} + a_{m+2} + \dots + a_{m+p} < d_m a_m/\alpha, \text{ a constant independent of } p.$$

The sum $a_{m+1} + a_{m+2} + \dots + a_{m+p}$ increases as p increases, but is always less than the constant $d_m a_m/\alpha$; this sum therefore tends to a limit when p tends to ∞ and therefore the series Σa_n is convergent.

Suppose next that $g_n < -\alpha < 0$ if $n \geqq m$. In this case $d_n a_n < d_{n+1} a_{n+1}$ if $n \geqq m$, and therefore

$$\frac{a_{m+p}}{a_m} = \frac{a_{m+1}}{a_m} \cdot \frac{a_{m+2}}{a_{m+1}} \cdots \frac{a_{m+p}}{a_{m+p-1}} > \frac{d_m}{d_{m+1}} \cdot \frac{d_{m+1}}{d_{m+2}} \cdots \frac{d_{m+p-1}}{d_{m+p}},$$

so that $$a_{m+p} > d_m a_m / d_{m+p}, \quad p = 1, 2, 3, \ldots .$$

Now $d_m a_m$ is independent of p and the series $\Sigma(1/d_n)$ diverges; hence also the series Σa_n diverges.

Note 1. If g_n tends to a limit l *which is not zero*, the sign of g_n for sufficiently large values of n will be that of l, and therefore Σa_n will converge or diverge according as l is positive or negative. Or again if the *minimum limit* of g_n is positive Σa_n will converge, while if the *maximum limit* of g_n is negative Σa_n will diverge.

Note 2. The proof for divergence shows that if a_{n+1}/a_n is greater than b_{n+1}/b_n for $n \geqq m$, the series Σa_n diverges if Σb_n diverges. It may be proved in the same way that if a_{n+1}/a_n is less than b_{n+1}/b_n for $n \geqq m$ the series Σa_n converges if Σb_n converges.

When a_{n+1}/a_n tends to unity the Test Ratio fails; Kummer's Theorem leads to a test for this case, usually called Raabe's Test, the terms a_n being all *positive*.

Raabe's Test. *The series Σa_n will converge or diverge according as*

$$n\left(\frac{a_n}{a_{n+1}} - 1\right) > 1 + \alpha > 1 \quad \text{or} \quad n\left(\frac{a_n}{a_{n+1}} - 1\right) < 1 - \alpha < 1;$$

when $n \geqq m$, a fixed integer (or, according as the limit for n tending to infinity of this expression is greater than 1 *or less than* 1).

In Kummer's Test let $d_n = n$; the series $\Sigma(1/n)$ diverges and therefore we have for convergence or divergence ($\alpha > 0$)

$$n\frac{a_n}{a_{n+1}} - (n+1) > \alpha \text{ or } < -\alpha;$$

that is, $$n\left(\frac{a_n}{a_{n+1}} - 1\right) > 1 + \alpha > 1 \text{ or } < 1 - \alpha < 1.$$

Gauss's Test. *Suppose that a_n/a_{n+1} can be expressed in the form*

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\mu}{n} + \frac{A_n}{n^\lambda}, \quad \begin{cases} \lambda > 1 \\ |A_n| < k, \text{ a constant for every } n. \end{cases}$$

The series Σa_n will converge if $\mu > 1$ but will diverge if $\mu \leqq 1$, the terms a_n being all positive.

Raabe's Test proves the theorem for $\mu>1$ and for $\mu<1$. For the case $\mu=1$, let $d_n=n\log n$; the series $\Sigma(1/d_n)$ diverges (Ex. 2 below). Now

$$g_n=d_n\frac{a_n}{a_{n+1}}-d_{n+1}=(n+1)\log\frac{n}{n+1}+\frac{A_n\log n}{n^{\lambda-1}}.$$

But
$$\mathop{\mathcal{L}}_{n\to\infty}\frac{\log n}{n^{\lambda-1}}=0 \text{ (§ 25, Ex. 6)}$$

and
$$\mathop{\mathcal{L}}_{n\to\infty}(n+1)\log\frac{n}{n+1}=\mathop{\mathcal{L}}_{n\to\infty}(n+1)\log\left(1-\frac{1}{n+1}\right)=-1,$$

and therefore g_n is negative when n is sufficiently large so that Σa_n diverges when $\mu=1$.

The test in Ex. 1 is often useful; the test in Ex. 4 is theoretically important.

Ex. 1. *Cauchy's Condensation Test.* If $\Sigma f(n)$ is a series of positive terms and if $f(n)>f(n+1)$, show that $\Sigma f(n)$ converges or diverges according as the series $\Sigma 2^n f(2^n)$ converges or diverges.

Of course the inequality $f(n)>f(n+1)$ need only begin when n is greater than some integer m, but there is no loss of generality in supposing it to hold from $n=1$. Proceed as in the case of the series $\Sigma(1/n^{a})$, (*E.T.* pp. 380, 381) and take the terms in groups of 2, 2^2, 2^3, ..., 2^n, ...

If $2^n \leqq \mu < 2^{n+1}$ we have

$$\begin{aligned}\sum_{n=2}^{\mu} f(n)&=[f(2)+f(3)]+[f(2^2)+f(5)+f(6)+f(7)]\\&\quad+[f(2^3)+f(9)+\ldots+f(15)]+\ldots+[f(2^n)+f(2^n+1)+\ldots+f(\mu)]\\&<2f(2)+2^2f(2^2)+2^3f(2^3)+\ldots+2^nf(2^n).\end{aligned}$$

When $n\to\infty$ so does μ, and therefore $\Sigma f(n)$ converges if $\Sigma 2^n f(2^n)$ converges.

By grouping as follows

$$[f(3)+f(2^2)]+[f(5)+\ldots+f(2^3)]+[f(9)+\ldots+f(2^4)]+\ldots$$

we see that these groups of terms are respectively greater than

$$\tfrac{1}{2}\,.\,2^2f(2^2),\quad \tfrac{1}{2}\,.\,2^3f(2^3),\quad \tfrac{1}{2}\,.\,2^4f(2^4),\ \ldots$$

so that $\Sigma f(n)$ diverges if $\Sigma 2^n f(2^n)$ diverges.

It is easy to prove that $\Sigma f(n)$ converges or diverges according as $\Sigma \mu^n f(\mu^n)$ converges or diverges where μ is any integer not less than 2.

In § 148, Ex. 6, it is proved that μ may be taken to be e. (See Chrystal's *Algebra*, Part II. p. 124.)

Ex. 2. The series $\sum \frac{1}{n(\log n)^\alpha}$, $n > 1$, converges or diverges according as $\alpha > 1$ or $\alpha \leqq 1$.

Here $f(n) = \frac{1}{n(\log n)^\alpha}$; $2^n f(2^n) = \frac{2^n}{2^n(\log 2^n)^\alpha} = \frac{1}{n^\alpha(\log 2)^\alpha}$.

But (*E.T.* pp. 380, 381) $\Sigma(1/n^\alpha)$ converges or diverges according as $a > 1$ or $a \leqq 1$; so therefore does the given series.

Ex. 3. **The Hypergeometric series.** The following series

$$1 + \frac{\alpha . \beta}{1 . \gamma} x + \frac{\alpha(\alpha+1) . \beta(\beta+1)}{1 . 2 . \gamma(\gamma+1)} x^2 + \frac{\alpha(\alpha+1)(\alpha+2) . \beta(\beta+1)(\beta+2)}{1 . 2 . 3 . \gamma(\gamma+1)(\gamma+2)} x^3$$
$$+ \ldots + \frac{\alpha(\alpha+1)\ldots(\alpha+n-1) . \beta(\beta+1)\ldots(\beta+n-1)}{1 . 2 . \ldots n . \gamma(\gamma+1)\ldots(\gamma+n-1)} x^n + \ldots$$

is called the *Hypergeometric Series* and is usually denoted by the symbol $F(\alpha, \beta, \gamma, x)$. The numbers α, β, γ, x are called the *elements* of the series and x alone is here considered as a variable, the elements α, β, γ being taken to be constants.

The series is symmetric in α and β so that $F(\alpha, \beta, \gamma, x) = F(\beta, \alpha, \gamma, x)$, and if either α or β is a negative integer, the series terminates. The element γ must not be a negative integer because, after a certain stage, each term of the series would have a zero denominator.

Take the term in x^n as a_n; then

$$\frac{a_{n+1}}{a_n} = \frac{(\alpha+n)(\beta+n)}{(n+1)(\gamma+n)} x \to x \text{ when } n \to \infty$$

so that the series converges absolutely if $|x| < 1$ and diverges if $|x| > 1$.

If $x = 1$ we find that

$$\frac{a_n}{a_{n+1}} = 1 + \frac{1+\gamma-\alpha-\beta}{n} + \frac{A_n}{n^2}$$

where $A_n \to \alpha^2 + \alpha\beta + \beta^2 - (\gamma+1)(\alpha+\beta) + \gamma$ when $n \to \infty$. When n is large all the terms are of the same sign since a_n/a_{n+1} differs little from 1 for large values of n. Gauss's Test may therefore be applied since $|A_n|$ is finite for every value of n. Hence the Hypergeometric Series, when $x = 1$, converges if $1 + \gamma - \alpha - \beta > 1$, that is, if $\gamma > \alpha + \beta$ and diverges if $\gamma \leqq \alpha + \beta$.

The student should study this example carefully as the Hypergeometric Series is of very great importance.

Ex. 4. The series Σa_n of positive terms converges or diverges according as the maximum limit G of $a_n^{1/n}$ is less than or greater than unity.

(i) Suppose $G < r < 1$. There is therefore at most a *finite* number of values of $a_n^{1/n}$ which exceed r; let all such values of $a_n^{1/n}$ be included in the first m values. Hence $a_n^{1/n} \leqq r$, that is, $a_n \leqq r^n$, if $n > m$, and therefore the series converges.

(ii) Suppose $G > r > 1$. There is therefore in this case an infinite number of values of a_n greater than r^n so that the series must diverge.

This test is often called **Cauchy's Test** and, though not so useful in ordinary applications as the Ratio Test (which is often called **d'Alembert's Test**), is of great theoretical importance. See Bromwich, *Inf. Ser.* (2nd Ed.), p. 32.

Ex. 5. Show that the power series $\Sigma a_n x^n$ converges or diverges according as $|x| < R$ or $|x| > R$ where $1/R$ is the maximum limit of $|a_n|^{1/n}$.

61. Tests of Abel and Dirichlet. An inequality that is very useful in the discussion of convergence is given in the Lemma (*E.T.* p. 451), known as Abel's Inequality, namely: if (c_n) is a decreasing sequence of positive numbers and if, for $r \leqq n$,

$$A > u_1 + u_2 + \ldots + u_r > B$$

where A and B are constants

$$c_1 A > c_1 u_1 + c_2 u_2 + \ldots + c_n u_n > c_1 B.$$

If $A > 0$ and $B = -A$ the inequality may be expressed as

$$|c_1 u_1 + c_2 u_2 + \ldots + c_n u_n| < c_1 A.$$

In the following tests the terms of Σu_n need not be all of the same sign.

Abel's Test. *A convergent (not necessarily absolutely convergent) series Σu_n remains convergent if each of its terms $u_1, u_2, u_3, \ldots$ is multiplied by a factor $a_1, a_2, a_3, \ldots$ provided the sequence (a_n) is monotonic and $|a_n|$ is less than a constant k for every n.*

The sequence (a_n), being monotonic and bounded, converges to a limit, a say. Let $c_n = a - a_n$ if (a_n) is an increasing sequence but $c_n = a_n - a$ if (a_n) is a decreasing sequence; the sequence (c_n) is therefore a decreasing sequence of positive numbers which has zero as its limit.

Now $a_n u_n = a u_n - c_n u_n$ or $a_n u_n = a u_n + c_n u_n$ according as $c_n = a - a_n$ or $c_n = a_n - a$; since Σu_n converges it is sufficient to prove that $\Sigma c_n u_n$ converges.

Let ${}_pR_n = c_{n+1} u_{n+1} + c_{n+2} u_{n+2} + \ldots + c_{n+p} u_{n+p}$. The series Σu_n is convergent and therefore there are constants A and B such that

$$A > u_{n+1} + u_{n+2} + \ldots + u_{n+p} > B, \quad p = 1, 2, 3, \ldots.$$

Hence, by Abel's Inequality,

$$A c_{n+1} > {}_pR_n > B c_{n+1}.$$

But $c_{n+1} \to 0$ when $n \to \infty$ and therefore ${}_pR_n \to 0$ when $n \to \infty$

whatever integer p may be so that $\Sigma c_n u_n$ and therefore also $\Sigma a_n u_n$ is convergent. (For the notation ${}_pR_n$ see *E.T.* p. 379.)

Dirichlet's Test. *If the series Σu_n oscillates finitely and if (c_n) is a decreasing sequence of positive numbers which has zero as its limit, the series $\Sigma c_n u_n$ is convergent.*

Let $${}_pR_n = u_{n+1} + u_{n+2} + \ldots + u_{n+p};$$
then, since Σu_n oscillates finitely, there are constants A and B such that $A > {}_pR_n > B$ for every value of n and p. If the rest of the notation is the same as in the proof of Abel's Test we have

$${}_pR_n = c_{n+1}u_{n+1} + c_{n+2}u_{n+2} + \ldots + c_{n+p}u_{n+p}.$$
$$Ac_{n+1} > {}_pR_n > Bc_{n+1},$$

and therefore ${}_pR_n \to 0$ when $n \to \infty$ for every integral value of p since $c_{n+1} \to 0$ when $n \to \infty$. Hence $\Sigma c_n u_n$ converges.

Ex. 1. If θ is neither zero nor a multiple of 2π the series

$$\sum_1^\infty \frac{\cos n\theta}{n}, \quad \sum_1^\infty \frac{\sin n\theta}{n}$$

are convergent. (The second series is zero if $\theta = 0, \pm\pi, \pm 2\pi, \ldots$.)

Ex. 2. Discuss the convergence of the series

$$\sum_1^\infty (-1)^{n-1}\frac{\cos n\theta}{n}, \quad \sum_1^\infty (-1)^{n-1}\frac{\sin n\theta}{n}.$$

62. Uniform Convergence. When the terms of a series are functions of a variable x, each term $u_n(x)$ being defined for the range $a \leqq x \leqq b$, that is, for the closed interval (a, b), the sum of the series when convergent will be a function $S(x)$ of x. For any given, or fixed, value of x in the interval (a, b) the condition for convergence is, if ${}_pR_n(x)$ denote the partial remainder after n terms (*E.T.* p. 379) and $S_n(x)$ the sum of the first n terms,

$$|{}_pR_n(x)| = |S_{n+p}(x) - S_n(x)| < \varepsilon \text{ if } n \geqq m,\ p = 1, 2, 3, \ldots \quad (1)$$

When x changes so, as a rule, will the integer m; if m is such that the inequality (1) is true whatever value x may have in the closed interval (a, b) the series is said to converge **uniformly** with respect to x in the closed interval.

Instead of the partial remainder ${}_pR_n(x)$ we may take the complete remainder $R_n(x)$; the two forms of the condition are equivalent, that is, given one of the conditions the other may be deduced from it.

If $|R_n(x)| < \varepsilon$ when $n \geqq m$ and $a \leqq x \leqq b$, then, $p = 1, 2, 3, \ldots$

$$|{}_pR_n(x)| = |R_n(x) - R_{n+p}(x)| \leqq |R_n(x)| + |R_{n+p}(x)| < 2\varepsilon.$$

If $|{}_pR_n(x)| < \varepsilon$ when $n \geqq m$, $p = 1, 2, 3, \ldots$ and $a \leqq x \leqq b$, then

$$|R_n(x)| = \underset{p \to \infty}{\mathcal{L}} |{}_pR_n(x)| \leqq \varepsilon < 2\varepsilon.$$

On pages 385, 386 of the *Elementary Treatise* some important theorems are proved. Theorem III, on p. 386, may in substance be stated in the following form and this test of uniform convergence is frequently cited as the "M-Test" or "Weierstrass's M-Test."

The M-Test. *If each term of the series* $\Sigma u_n(x)$ *is defined for the range* $a \leqq x \leqq b$ *and if, for each term,* $|u_n(x)| < M_n$, a number independent of x, *the series* $\Sigma u_n(x)$ *converges uniformly for the range* $a \leqq x \leqq b$ *provided the series* ΣM_n *is convergent.*

Obviously
$$\sum^{m+p} u_r(x) < \sum_{n=m}^{m+p} M_n,$$

and if the second sum is less than ε the choice of m that makes it so does not depend upon x so that the inequality (1) is satisfied for the range $a \leqq x \leqq b$.

The Tests of Abel and Dirichlet are easily adapted so as to be tests for uniform convergence; the following statements are from Bromwich (*l.c.* p. 125).

Abel's Test. *The series* $\Sigma a_n(x)u_n(x)$ *converges uniformly in the closed interval* (a, b) *if the following conditions are fulfilled:* (i) $\Sigma u_n(x)$ *converges uniformly in* (a, b); (ii) $a_n(x)$ *is, for a fixed value of* x *in* (a, b), *positive and does not increase as* n *increases; and* (iii) $a_1(x) < k$, *a constant, for the range* $a \leqq x \leqq b$.

By condition (i) if $a \leqq x \leqq b$,

$$|u_{n+1}(x) + u_{n+2}(x) + \ldots + u_{n+p}(x)| < \varepsilon \text{ if } n \geqq m,$$

and therefore, by Abel's Inequality, with $A = \varepsilon$, $B = -\varepsilon$,

$$\sum_{r=n+1}^{n+p} a_r(x)u_r(x) < \varepsilon a_{n+1}(x) < \varepsilon k \text{ if } n \geqq m,\ a \leqq x \leqq b,$$

since $a_{n+1}(x) \leqq a_1(x)$ and $a_1(x) < k$. Hence $\Sigma a_n(x)u_n(x)$ converges uniformly in (a, b).

Special cases of this theorem arise if $a_n(x)$ is independent of x or if $u_n(x)$ is independent of x.

Dirichlet's Test. *The series $\Sigma a_n(x)u_n(x)$ converges uniformly in the closed interval (a, b) if the following conditions are fulfilled :*

(i) *$\Sigma u_n(x)$ oscillates, its values lying between $-k$ and k, where k is a fixed constant ;* (ii) *$a_n(x)$ is, for a fixed value of x, positive and does not increase as n increases ; and* (iii) *$a_n(x)$, when $n \to \infty$, tends* **uniformly** *to zero for all values of x in the closed interval (a, b).*

By condition (i) if $a \leqq x \leqq b$,

$$| u_{n+1}(x) + u_{n+2}(x) + \dots + u_{n+p}(x) | < 2k$$

and by condition (iii) m can be chosen so that $a_n(x) < \varepsilon$ if $n \geqq m$ and $a \leqq x \leqq b$. Hence by Abel's Inequality

$$\sum_{r=n+1}^{n+p} a_r(x)u_r(x) \; < 2\epsilon k \text{ if } n \geqq m, \; a \leqq x \leqq b.$$

Particular cases arise if $u_n(x)$ is independent of x while Σu_n either converges or oscillates finitely or if $a_n(x)$ is independent of x.

Ex. 1. The series $\Sigma x^n/n!$ converges uniformly in the interval $(-a, a)$, where a is arbitrarily large.

Let a be any positive number, however large, and let $M_n = a^n/n!$. The M-Test shows that the given series converges uniformly if $|x| \leqq a$. The series is often said to converge uniformly for every value of x, or "for every x."

Ex. 2. The series Σn^{-x} and $\Sigma(\log n) \,.\, n^{-x}$ converge uniformly if $x \geqq 1 + k > 1$.

For the first series let $M_n = 1/n^{1+k}$ and the M-Test applies.

The second series is obtained by differentiating the first. Now $(\log n) \,.\, n^{-\alpha} \to 0$ when $n \to \infty$ if $\alpha > 0$. Therefore there is an integer m such that $0 < (\log n) \,.\, n^{-\alpha} < C$, a constant, if $n \geqq m$. Now let $\alpha = \frac{1}{2}k$, and we have, if $x \geqq 1 + k > 1$,

$$\frac{\log n}{n^x} \leqq \frac{\log n}{n^{1+k}} < C \,.\, \frac{1}{n^{1+\frac{1}{2}k}}, \; n \geqq m.$$

If $M_n = C/n^{1+\frac{1}{2}k}$ the M-Test applies.

Ex. 3. If Σc_n converges the series $\Sigma(c_n/n^x)$ converges uniformly in the closed interval $(0, 1)$.

In Abel's Test, let $a_n(x) = n^{-x}$ and $u_n(x) = c_n$, independent of x.

Ex. 4. The series $\sum \frac{\quad}{n^a}$, $\sum \frac{\quad}{n^a}$ converge uniformly for every x if $\alpha > 1$, but if $0 < \alpha \leqq 1$ they converge uniformly for the range $0 < \theta \leqq x \leqq 2\pi - \theta$.

If $\alpha > 1$ apply the M-Test ; if $0 < \alpha \leqq 1$ apply Dirichlet's Test, taking $u_n(x)$ equal to $\cos nx$ and to $\sin nx$ respectively.

63. Tannery's Theorem. The following theorem, stated by Tannery in his *Fonctions d'une Variable*, 2nd Ed., § 183, and called by Bromwich "Tannery's Theorem," is closely related to the M-Test for uniform convergence and is, in fact, proved by Tannery as a particular case of that test.

Let
$$F(n)=\sum_{r=0}^{N} u_r(n) \quad \text{.................................(1)}$$
where $u_r(n)$ is a function of n, and N is also a function of n that tends to infinity with n. If $u_r(n)$, when r is fixed, tends to a limit, v_r say, when $n \to \infty$, will the sum $F(n)$ tend to $\sum_0^\infty v_r$ when $n \to \infty$?

An example is given by taking $F(n)$ equal to $(1+1/n)^n$ and deducing the value of e as a series ($E.T.$ § 48). The problem occurs with sufficient frequency to justify the statement of a general theorem that will save repetitions.

Tannery's Theorem. *Suppose that the following conditions are satisfied:*

(i) $\underset{n\to\infty}{\mathcal{L}} u_r(n)=v_r$, *when r is fixed;*

(ii) $|u_r(n)| \leqq M_r$, *where M_r is independent of n;*

(iii) $\sum_0^\infty M_r$ *is convergent.*

When these conditions are satisfied $F(n) \to \sum_0^\infty v_r$ when $n \to \infty$.

By (i) and (ii) $|v_r| \leqq M_r$ and therefore by (iii) Σv_r converges.

Again, since n and therefore N is to tend to infinity we may always suppose that N is larger than m, whatever integer m may be, and we may therefore express $F(n)-\Sigma v_r$ in the form
$$F(n)-\sum^{\infty} v_r=\sum_{r=0}^{m}[u_r(n)-v_r]+\sum_{r=m+1}^{N} u_r(n)-\sum_{r=m+1}^{\infty} v_r. \quad \text{.....(2)}$$
Now ΣM_r converges and therefore m may be chosen so that
$$\left|\sum_{r=m+1}^{\infty} v_r\right| \leqq \sum_{r=m+1}^{\infty} M_r<\varepsilon, \quad \left|\sum_{r=m+1}^{N} u_r(n)\right| \leqq \sum_{r=m+1}^{\infty} M_r<\varepsilon \quad \text{.(3)}$$
where ε has the usual meaning.

The value of m in the equalities (3) depends only on the series ΣM_r and is therefore independent of n; when m has been chosen so as to satisfy the inequalities (3) let it be kept fast. The first sum on the right of equation (2) contains a finite number of terms and therefore by condition (i) the number n_1

can be chosen so that, if $n>n_1$, the first sum on the right of (2) will be numerically less than ε. Hence if $n>n_1$

$$F(n)-\sum_{r=0}^{\infty}v_r\,\Big|\leqq\Big|\sum_{r=0}^{m}[u_r(n)-v_r \quad + \sum_{r=m+1} u_r(n) \qquad \sum_{r=m+1}^{\infty} v_r$$

that is, $$F(n)-\sum_{r=0}^{\infty}v_r \;<3\varepsilon \text{ if } n>n_1$$

and therefore $$\mathop{\mathcal{L}} F(n)=\sum_{r=0}^{\infty}v_r.$$

Ex. If $F(n)=\left(1+\frac{x}{n}\right)^n$ show that the limit of $F(n)$ when n tends to infinity is the series

$$\sum_{r=0}^{\infty}\frac{x^r}{r!}.$$

Here $u_r(n)=\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{r-1}{n}\right)\cdot\frac{x^r}{r!}.$

$$N=n,\ v_r=\frac{x^r}{r!},\ M_r=\frac{a^r}{r!}$$

where a is any fixed positive number.

64. Abel's Theorem. When the interval of convergence (*E.T.* p. 384) of a power series is $(-R, R)$ the interval may be changed to $(-1, 1)$ by substituting x/R for x; we therefore suppose that in this and the next article each power series converges for $|x|<1$ and diverges for $|x|>1$. The behaviour of the series when $x\to 1$ or when $x\to -1$ from within the interval will now be considered; *for definiteness x will be supposed to tend to* $+1$.

Let $$f(x)=\sum_{n=0}^{\infty}a_nx^n,\ s_n=a_0+a_1+\dots+a_n,$$

the series for $f(x)$ converging for $|x|<1$.

Abel's Theorem. (i) *If $s_n\to s$ (a finite number) when $n\to\infty$ the function $f(x)\to s$ when $x\to 1$*; (ii) *if $s_n\to\infty$ (or to $-\infty$) when $n\to\infty$ the function $f(x)\to\infty$ (or to $-\infty$) when $x\to 1$.*

If $$0<x<1,\quad 1/(1-x)=\sum_{0}^{\infty}x^n,$$

and therefore, if the series for $f(x)$ and $1/(1-x)$ are multiplied,

$$f(x)/(1-x)=\sum_{n=0}^{\infty}s_nx^n=\sum_{n=0}^{m}s_nx^n+\sum_{n=m+1}^{\infty}s_nx^n$$

where the series Σs_nx^n converges for $|x|<1$.

(i) Let $s_n \to s$. It is then possible to choose m so that, given ε,

$$s-\varepsilon < s_n < s+\varepsilon \text{ if } n \geqq m,$$

and when m is thus chosen it is to be kept fixed. Next let

$$f_m(x)=\sum_{n=0}^{m} s_n x^n, \quad F_m(x)=\sum_{n=m+1}^{\infty} s_n x^n,$$

and therefore $f(x)=(1-x)f_m(x)+(1-x)F_m(x)$;

then $f_m(x)$ is finite if $x\to 1$ so that $(1-x)f_m(x)\to 0$ if $x\to 1$.

Again, $\quad F_m(x) < (s+\varepsilon)\sum_{m+1}^{\infty} x^n$ but $> (s-\varepsilon)\sum_{m+1}^{\infty} x^n$,

that is, $\quad F_m(x) < (s+\varepsilon)x^{m+1}/(1-x)$ but $> (s-\varepsilon)x^{m+1}/(1-x)$,

and therefore $\quad f(x) < (1-x)f_m(x)+(s+\varepsilon)x^{m+1}$

but $\quad f(x) > (1-x)f_m(x)+(s-\varepsilon)x^{m+1}$.

Hence $$\mathop{\mathcal{L}}_{x\to 1} f(x) \leqq s+\varepsilon \text{ but } \geqq s-\varepsilon.$$

and as ε is arbitrarily small $f(x)\to s$ when $x\to 1$; therefore

$$\mathop{\mathcal{L}}_{x\to 1}\left(\sum_{n=0}^{\infty} a_n x^n\right)=\sum_{n=0}^{\infty} a_n=\sum_{n=0}^{\infty}\left(\mathop{\mathcal{L}}_{x\to 1} a_n x^n\right).$$

(ii) Let $s_n\to\infty$ so that the series Σa_n is divergent; if $s_n\to-\infty$ the sign of every term in $\Sigma a_n x^n$ may be changed so that there is no loss of generality in this restriction on the limit of s_n.

Let G be any given arbitrarily large positive number; m may be chosen so that $s_n > G' > G$ if $n \geqq m$. In this case, x being positive and less than unity,

$$F_m(x)=\sum_{n=m+1}^{\infty} s_n x^n > G'\sum_{m+1}^{\infty} x^n = G' x^{m+1}/(1-x),$$

and therefore, since $(1-x)f_m(x)\to 0$ when $x\to 1$, $f_m(1)$ being finite,

$$\mathop{\mathcal{L}}_{x\to 1} f(x) \geqq G' > G.$$

As G is arbitrarily large, $f(x)\to\infty$ when $x\to 1$.

Ex. 1. $\log(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3-\frac{1}{4}x^4+\dots$, $|x|<1$.

The series converges for $x=1$; the series for $\log 2$ is therefore obtained by putting $x=1$. The series diverges for $x=-1$ and $\log(1+x)\to-\infty$ when $x\to-1$.

Ex. 2. If the series $\Sigma a_n x^n$ and $\Sigma b_n x^n$ converge when $|x|<1$ their product is (*E.T.* p. 388) $\Sigma c_n x^n$ where

$$c_n = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \ldots + a_n b_0.$$

Show that if the series Σc_n converges the product of the convergent series Σa_n and Σb_n is Σc_n whether Σa_n and Σb_n converge absolutely or conditionally.

The statement follows at once from Abel's Theorem because the series tend to Σa_n, Σb_n and Σc_n respectively when $x \to 1$.

65. Cesàro's Theorem. This theorem deals with the limit for $x \to 1$ of the ratio of two power series which diverge when $x \to 1$.

Let $f(x) = \Sigma a_n x^n$, $g(x) = \Sigma b_n x^n$ where the series Σa_n and Σb_n are both divergent; we suppose that each diverges to $+\infty$, as in case (ii) of Abel's Theorem, and that the coefficients b_n are all *positive* for $n > n'$, some given number.

There are two cases.

(i) *If the quotient a_n/b_n tends to a (finite) limit l when $n \to \infty$ the quotient $f(x)/g(x)$ will tend to l when $x \to 1$.*

We can choose m so that $l - \varepsilon < a_n/b_n < l + \varepsilon$ if $n \geqq m$ and therefore, since $b_n > 0$ and $x^n > 0$,

$$(l-\varepsilon) \sum_{n=m+1}^{\infty} b_n x^n < \sum_{n=m+1}^{\infty} a_n x^n < (l+\varepsilon) \sum_{n=m+1}^{\infty} b_n x^n.$$

Let
$$\sum_{n=0}^{m} a_n x^n = \varphi_m(x), \quad \sum_{n=m+1}^{\infty} a_n x^n = \Phi_m(x),$$
$$\sum_{0}^{m} b_n x^n = \psi_m(x), \quad \sum_{n=m+1}^{\infty} b_n x^n = \Psi_m(x);$$

then
$$(l-\varepsilon)\Psi_m(x) < \Phi_m(x) < (l+\varepsilon)\Psi_m(x).$$

Next
$$\frac{\Phi_m(x)}{\Psi_m(x)} = \frac{f(x) - \varphi_m(x)}{g(x) - \psi_m(x)} = \frac{f(x)}{g(x)} \cdot \frac{1 - \varphi_m(x)/f(x)}{1 - \psi_m(x)/g(x)},$$

so that
$$\frac{f(x)}{g(x)} = \frac{\Phi_m(x)}{\Psi_m(x)} \cdot \frac{1 - \psi_m(x)/g(x)}{1 - \varphi_m(x)/f(x)}.$$

When $x \to 1$ both $f(x)$ and $g(x)$ tend to $+\infty$ (by Abel's Theorem) while $\varphi_m(x)$ and $\psi_m(x)$, being each the sum of a finite number of terms, remain finite; the fraction

$$\{1 - \psi_m(x)/g(x)\} \div \{1 - \varphi_m(x)/f(x)\}$$

therefore tends to 1 when $x \to 1$. Again the fraction $\Phi_m(x)/\Psi_m(x)$ cannot when $x \to 1$ fall outside the interval $(l-\varepsilon, l+\varepsilon)$. Hence since ε is arbitrarily small, $f(x)/g(x) \to l$ when $x \to 1$.

(ii) *If the quotient a_n/b_n tends to ∞ when $n \to \infty$, the quotient $f(x)/g(x)$ will tend to ∞ when $x \to 1$.*

We can choose m so that $a_n/b_n > G' > G$ for $n \geqq m$, where G is any given arbitrarily large positive number, and therefore $\Phi_m(x)/\Psi_m(x) > G'$. As before, the limit of $f(x)/g(x)$ when $x \to 1$ is seen to be greater than G and therefore $f(x)/g(x)$ tends to ∞ when $x \to 1$.

For the proof compare that of Theorem II, p. 420, of the *Elementary Treatise.*

Ex. 1. $\underset{x \to 1}{\mathcal{L}} (1-x)^{\frac{1}{2}}(x + x^4 + x^9 + \ldots + x^{n^2} + \ldots) = \frac{1}{2}\sqrt{\pi}.$

The symbol $[\sqrt{n}]$ is used to denote the greatest integer contained in $\sqrt{n}$. For example,

$$[\sqrt{2}] = 1\ ;\ [\sqrt{4}] = 2\ ;\ [\sqrt{7}] = 2\ ;\ [\sqrt{29}] = 5, \ldots$$

Let $\qquad \varphi(x) = x + x^4 + x^9 + \ldots + x^{n^2} + \ldots$

and $\qquad f(x) = [\sqrt{1}]x + [\sqrt{2}]x^2 + \ldots + [\sqrt{n}]x^n + \ldots\ ;$

then $\qquad (1-x)f(x) = \varphi(x)\ ;\ (1-x)^{\frac{1}{2}}\varphi(x) = \dfrac{f(x)}{(1-x)^{-\frac{3}{2}}}.$

Now $\qquad (1-x)^{-\frac{3}{2}} = \displaystyle\sum_{n=0}^{\infty} \frac{3 \cdot 5 \cdot 7 \ldots (2n+1)}{2 \cdot 4 \cdot 6 \ldots (2n)} x^n = \Sigma b_n x^n,$ say,

and the series for $f(x)$ and $(1-x)^{-\frac{3}{2}}$ diverge when $x \to 1$. Hence, by Cesàro's Theorem,

$$\underset{x \to 1}{\mathcal{L}} (1-x)^{\frac{1}{2}}\varphi(x) = \underset{x \to 1}{\mathcal{L}} \frac{f(x)}{(1-x)^{-\frac{3}{2}}} = \underset{n \to \infty}{\mathcal{L}} \frac{[\sqrt{n}]}{b_n}$$

provided the limit for $n \to \infty$ exists.

From Exercises II. 29, $\quad b_n = (2n+1)a_n/\sqrt{n}$ where $a_n \to \pi^{-\frac{1}{2}}$ when $n \to \infty$ and therefore,

$$\mathcal{L} \frac{[\sqrt{n}]}{b_n} = \tfrac{1}{2}\pi^{\frac{1}{2}}.$$

The above proof is that of Cesàro ; for another proof see Bromwich, *Infinite Series* (2nd Ed.), p. 150, Ex. 4.

Ex. 2. If $f(x) = \Sigma a_n x^n$, $\ s_n = a_0 + a_1 + a_2 + \ldots + a_n$, and if

$$(s_0 + s_1 + s_2 + \ldots + s_n)/(n+1)$$

tends to l when $n \to \infty$, prove that $f(x)$ also tends to l when $x \to 1$.

(Frobenius.)

Let $s_0 + s_1 + s_2 + \ldots + s_n = t_n$; th (see the proof of Abel's Theorem (i)),

$$f(x) = \Sigma a_n x^n = \frac{\Sigma s_n x^n}{(1-x)^{-1}} = \frac{\Sigma t_n x^n}{(1-x)^{-2}} = \frac{\Sigma t_n x^n}{\Sigma (n+1) x^n},$$

and therefore, by Cesàro's Theorem,

$$\mathop{\mathcal{L}}_{x\to 1} f(x) = \mathop{\mathcal{L}}_{n\to\infty} \frac{t_n}{n+1} = l.$$

If $s_n \to l$ the mean $t_n/(n+1)$ also tends to l, by Cauchy's Theorem (§ 20, Ex. 3) and in this case Frobenius's Theorem gives no more information than Abel's; but it is quite possible that the mean may tend to a limit though s_n does not, and Frobenius's Theorem does then give new information about the behaviour of $f(x)$ when $x\to 1$.

The process may obviously be carried further. Suppose that neither s_n nor the mean $t_n/(n+1)$ tends to a limit and take a second mean, namely $\sigma_n/\frac{1}{2}(n+1)(n+2)$ where

$$\begin{aligned}\sigma_n &= (n+1)s_0 + ns_1 + (n-1)s_2 + \ldots + 2s_{n-1} + s_n \\ &= t_0 + t_1 + t_2 + \ldots + t_n.\end{aligned}$$

If this means tends to l when $n\to\infty$ then $f(x)\to l$ when $x\to 1$. For

$$f(x)(1-x)^{-3} = \sum_0^\infty \sigma_n x^n\ ;\ (1-x)^{-3} = \sum_0^\infty \tfrac{1}{2}(n+1)(n+2)x^n,$$

and therefore
$$f(x) = \frac{\Sigma\sigma_n x^n}{\Sigma\frac{1}{2}(n+1)(n+2)x^n},$$

so that
$$\mathop{\mathcal{L}}_{x\to 1} f(x) = \mathop{\mathcal{L}}_{n\to\infty} \frac{\sigma_n}{\frac{1}{2}(n+1)(n+2)} = l.$$

by Cesàro's Theorem.

66. Derangement of a Series. It has been seen (§ 59) that when a series Σa_n is absolutely convergent no derangement of its terms affects its sum. Suppose now that ΣA_m is a convergent series whose sum is S and that each term A_m is itself an infinite series, say

$$A_m = \sum_{n=1}^\infty a_{m,n} = a_{m,1} + a_{m,2} + \ldots + a_{m,n} + \ldots \quad \ldots\ldots\ldots\ldots(1)$$

where $m = 1, 2, 3, \ldots$, and therefore

$$S = \sum_{m=1}^\infty A_m = \sum_{m=1}^\infty \left(\sum_{n=1}^\infty a_{m,n}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

If summation is made first with respect to m, so that for $n = 1, 2, 3, \ldots$ we find

$$B_n = \sum_{m=1}^\infty a_{m,n} = a_{1,n} + a_{2,n} + \ldots + a_{m,n} + \ldots, \quad \ldots\ldots\ldots\ldots(3)$$

and then, the new sum being denoted by S',

$$S' = \sum_{n=1}^\infty B_n = \sum_{n=1}^\infty \left(\sum_{m=1}^\infty a_{m,n}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots(4)$$

will the new series be convergent and, if so, will S and S' be equal?

The following are sufficient (not necessary) conditions that the series (4) should be convergent and have the same sum as the series (2):

(i) the series (1) converges *absolutely* for every m say

$$\alpha_m = |a_{m,1}| + |a_{m,2}| + \ldots + |a_{m,n}| + \ldots;$$

(ii) the series $\Sigma\alpha_m$ is convergent, its sum being σ, say.

When these conditions are satisfied the series (4) is convergent, $S' = S$ and the series (3) is *absolutely* convergent for every n.

The proof follows from the theorem of § 59; because the terms $a_{m,n}$ of the series (2), which may be called a *double series*, as containing a "doubly infinite" number of terms, may be arranged so to form a single series $b_1 + b_2 + b_3 + \ldots$. The arrangement may be effected in many ways, but there are two ways of special importance.

Let the terms $a_{m,n}$ be arranged in tabular form (T) so that in any one row m is constant and in any one column n is constant. A_m is the sum of the mth row while B_n is the sum of the nth column.

$$\begin{matrix} a_{1,1}, & a_{1,2}, & a_{1,3}, & a_{1,4}, & \ldots\ldots \\ a_{2,1}, & a_{2,2}, & a_{2,3}, & a_{2,4}, & \\ a_{3,1}, & a_{3,2}, & a_{3,3}, & a_{3,4}, & \end{matrix} \qquad \text{(T)}$$

Arrangement by diagonals. Take the terms for which $(m+n)$ is constant, taking successively the terms for which $(m+n)$ is equal to 2, 3, 4, ... and, for each group in which $(m+n)$ has the same value, arrange the terms in descending order with respect to m. We thus find

$$a_{1,1} \mid a_{2,1},\ a_{1,2} \mid a_{3,1},\ a_{2,2},\ a_{1,3} \mid a_{4,1},\ a_{3,2},\ a_{2,3},\ a_{1,4} \mid \ldots$$

Each group lies on a "diagonal" of the array (T).

Arrangement by squares. The terms common to the first m rows and the first m columns form a square array. In the arrangement by "squares" the terms are taken from the mth row and the mth column of the square array, beginning with the term $a_{m,1}$, going on to the term $a_{m,m}$ and ending with the term $a_{1,m}$. Thus we find the successive groups

$$a_{1,1} \mid a_{2,1},\ a_{2,2},\ a_{1,2} \mid a_{3,1},\ a_{3,2},\ a_{3,3},\ a_{2,3},\ a_{1,3} \mid \ldots.$$

It is clear that both methods give a single series in which each term occurs once, and only once, in the table (T), while each term in the table appears once, and only once, in the single series.

Now let $T_p = b_1 + b_2 + \ldots + b_p$ and $S_p = |b_1| + |b_2| + \ldots + |b_p|$ so that S_p is the sum of the moduli of the terms of T_p. It is possible to choose m so large that all the terms of the sum S_p occur in the sum $\alpha_1 + \alpha_2 + \ldots + \alpha_m$ which is less than σ;

therefore S_p, which increases as p increases but is less than σ for every value of p, tends to a limit when $p \to \infty$ so that the series Σb_p is absolutely convergent. But, by hypothesis, to every term b_p there corresponds one and only one term $a_{m,n}$ and conversely to every term $a_{m,n}$ there correspor s one, and only one, term b_p so that the series (2) and (4) are both derangements of Σb_p and therefore both converge and $S' = S = \Sigma b_p$.

Further, if $\beta_m = | a_{1,n} | + | a_{2,n} | + \dots + | a_n$ the sum β_n is less than $\Sigma \mid b_p \mid$, and therefore, by the usual reasoning, the series (3) is absolutely convergent.

Again, if $S_{m,n}$ is the sum of the terms common to the first m rows and the first n columns of the array (T) the sum $S_{m,m}$ is the sum T_p, if $p = m^2$, when the terms of the array are arranged in squares; $S_{m,m}$ is the sum when there are m terms in the "side" of a square and $S_{n,n}$ the sum when there are n terms in a side, and both $S_{m,m}$ and $S_{n,n}$ tend to S when m and n tend respectively to infinity. Now $S_{m,n}$ lies between $S_{m,m}$ and $S_{n,n}$ in the sense that the difference between $S_{m,n}$ and either $S_{m,m}$ or $S_{n,n}$ is a sum of terms $b_r, b_s, b_t, \dots$; but the sum $|b_r| + |b_s| + |b_t| + \dots$ tends to zero when m and n tend to infinity, and therefore when m and n tend in any way to infinity $S_{m,n}$ also tends to S.

We thus have the result that when conditions (i) and (ii) are satisfied $S_{m,n}$ tends to the same limit when m and n tend to infinity in the following three ways:

$$(a) \mathop{\mathcal{L}}_{m \to \infty} \left(\mathop{\mathcal{L}}_{n \to \infty} S_{m,n} \right) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n}$$

$$(b) \mathop{\mathcal{L}}_{n \to \infty} \left(\mathop{\mathcal{L}}_{m \to \infty} S_{m,n} \right) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_n$$

$$(c) \ \mathcal{L} \ S_{m,n} = \mathcal{L} \ \sum a_{m,n}$$

where in (c) m and n tend *independently* to ∞, that is, the only restriction on m and n is that each becomes and remains larger than any given integer N.

The series in (a) and (b) are said to be formed by "repeated summation"; one of the numbers m, n tends first to infinity

and then the other and the two repeated summations give the same sum. In case (c) the "double series" $\Sigma a_{m,n}$ is said to converge.

It is possible, when conditions (i) and (ii) are not satisfied, that the limit given by (c) may exist and yet the series A_m and B_n may not converge, but into such cases we do not enter. See Bromwich, *Infinite Series*, Chapter V.

Cor. Multiplication of Series. If $a_{m,n}=b_m c_n$ and if the series Σb_m and Σc_n are absolutely convergent with sums B and C respectively the terms in the mth row of the array (T) will be $b_m c_1, b_m c_2, \ldots, b_m c_n, \ldots$ and therefore

$$A_m = b_m(c_1+c_2+\ldots+c_n+\ldots)=b_m C,$$

and

$$S=\sum_{m=1}^{\infty} A_m = \left(\sum_{1}^{\infty} b_m\right)C = BC$$

because conditions (i) and (ii) are satisfied. If now the terms in (T) are arranged by diagonals we find

$$BC=S=b_1c_1+(b_2c_1+b_1c_2)+(b_3c_1+b_2c_2+b_1c_3)+\ldots$$

which is the usual rule for the multiplication of two absolutely convergent series.

It would be more symmetrical to take $a_{m,n}=b_m x^m \,.\, c_n x^n$ and to let m, n take the values $0, 1, 2, \ldots$; by this notation we should get the form given on p. 388 of the *Elementary Treatise*.

The following examples 1-4 are from Bromwich, p. 86.

Ex. 1. If $a_{m,n}=c_m c_n$ where the numbers are positive, the double series $\Sigma a_{m,n}$ converges if Σc_m converges (say to C).

Here
$$S_{m,m}=(c_1+c_2+\ldots+c_m)^2 < C^2,$$
so that $S_{m,m}$ and therefore also $S_{m,n}$ tends to a limit.

Ex. 2. If $m+n=p$ and if $a_{m,n}=c_p/p$, the terms being all positive, the double series $\Sigma a_{m,n}$ converges if Σc_p converges (say to C).

Here, if we take the arrangement by diagonals, we have for the single series
$$\tfrac{1}{2}c_2+2(\tfrac{1}{3}c_3)+3(\tfrac{1}{4}c_4)+\ldots+(p-1)\left(\frac{c_p}{p}\right)+\ldots < C,$$
and the result follows. It may be seen similarly that if $a_{m,n}=d_p/p$ and if Σd_p diverges so does $\Sigma a_{m,n}$.

Ex. 3. The double series $\Sigma(m+n)^{-\alpha}$ converges if $\alpha > 2$ and diverges if $\alpha \leqq 2$, while $\Sigma m^{-\alpha} n^{-\beta}$ converges if $\alpha > 1$, $\beta > 1$.

Ex. 4. If a, c are positive (and $ac > b^2$ if $b < 0$) the series
$$\sum(am^2+2bmn+cn^2)^{-\lambda}$$
converges if $\lambda > 1$ and diverges if $\lambda \leqq 1$.

When A is the greatest of the three numbers a, c, $|b|$, we have
$$A(m+n)^2 > am^2+2bmn+cn^2 > 2[b+\surd(ac)]mn$$
and the result follows from Ex. 3.

Ex. 5. Show that, if $|x|=\xi$, $|y|=\eta$ and $2\xi\eta+\eta^2<1$,
$$(1-2xy+y^2)^{-\frac{1}{2}} = 1+\sum_{n=1}^{\infty} P_n(x)y^n$$
where $P_n(x)$ is a polynomial in x of the nth degree.

If $|y(2x-y)| < 1$ the Binomial Expansion gives
$$(1-2xy+y^2)^{-\frac{1}{2}} = 1+\sum_{n=1}^{\infty}\frac{1.3.5\ldots(2n-1)}{2.4.6\ldots(2n)}y^n(2x-y)^n \quad\ldots\ldots\ldots(1)$$
Let $|x|=\xi$ and $|y|=\eta$; then, if $2\xi\eta+\eta^2<1$,
$$(1-2\xi\eta-\eta^2)^{-\frac{1}{2}} = 1+\sum_{n=1}^{\infty}\frac{1.3.5\ldots(2n-1)}{2.4.6\ldots(2n)}\eta^n(2\xi+\eta)^n \quad\ldots\ldots\ldots(2)$$

Now the series in (2) is what the series in (1) becomes when *every term in it is made positive*, and as the series in (1) when thus treated is convergent (by (2)) its terms may be deranged and rearranged in powers of y. When so rearranged the coefficient of y^n is $P_n(x)$, the polynomial of the nth degree,
$$\frac{1.3.5\ldots(2n-1)}{1.2.3\ldots n}\left\{x^n-\frac{n(n-1)}{2(2n-1)}x^{n-2}+\frac{n(n-1)(n-2)(n-3)}{2.4.(2n-1)(2n-3)}x^{n-4}-\ldots\right\}.$$

Now $2\xi\eta+\eta^2<1$ if $\eta<(1+\xi^2)^{\frac{1}{2}}-\xi$ and this inequality is satisfied if $\xi\leqq 1$ and $\eta<\surd 2-1=0{\cdot}414$. Hence the series
$$1+\sum_{n=1}^{\infty}P_n(x)y^n$$
converges absolutely and uniformly with respect to x and absolutely and uniformly with respect to y for the ranges
$$|x|\leqq 1 \text{ and } |y|\leqq c<\surd 2-1.$$

The polynomials $P_n(x)$ are called *Legendre's Polynomials of degree n* or *Legendre's Coefficients of degree n* or *Zonal Harmonics of degree n*. (See, for example, MacRobert's *Spherical Harmonics*, Chapters IV, V.)

Ex. 6. Show that $P_n(-x)=(-1)^nP_n(x)$.

From the value of $P_n(x)$ in Ex. 5 it is obvious that $P_n(x)$ contains only even powers of x when n is even and only odd powers of x when n is odd; the result then follows. The relation may, however, be proved independently by expressing $(1+2xy+y^2)^{-\frac{1}{2}}$ in the two forms
$$[1-2(-x)y+y^2]^{-\frac{1}{2}} \text{ and } [1-2x(-y)+(-y)^2]^{-\frac{1}{2}},$$
which give the identical equation
$$1+\sum_{n=1}^{\infty}P_n(-x)y^n = 1+\sum_{n=1}^{\infty}P_n(x)(-y)^n$$
and therefore
$$P_n(-x)=(-1)^nP_n(x).$$

Ex. 7. Prove the following values where, for symmetry, $P_0(x)=1$.

$$P_0(x)=1\,;\ P_2(x)=\tfrac{3}{2}x^2-\tfrac{1}{2}\,;\ P_4(x)=\frac{5\,.\,7}{2\,.\,4}x^4-\frac{3\,.\,5}{2\,.\,4}2x^2+\frac{1\,.\,3}{2\,.\,4}\,;$$

$$P_1(x)=x\,;\ P_3(x)=\tfrac{5}{2}x^3-\tfrac{3}{2}x\,;\ P_5(x)=\frac{7\,.\,9}{2\,.\,4}x^5-\frac{5\,.\,7}{2\,.\,4}2x^3+\frac{3\,.\,5}{2\,.\,4}x.$$

EXERCISES VII.

1. Prove the conditions (*E.T.* p. 395) for the convergence or divergence of the Binomial Expansion of $(1+x)^m$ when $x=\pm 1$.

2. If $u_n=(n!)^2x^n/(2n)!$ the series Σu_n converges or diverges according as $|x|$ is less than 4 or greater than 4.

3. If $u_n=1\,.\,3\,.\,5\,...\,(2n-1)/2\,.\,4\,.\,6\,...\,2n$ the series Σu_n diverges.

4. The series $\dfrac{a}{b}+\dfrac{a(a+1)}{b(b+1)}+\dfrac{a(a+1)(a+2)}{b(b+1)(b+2)}+...$ converges or diverges according as $b-a$ (or the real part of $b-a$) is greater than 1 or not greater than 1.

5. If $0<x\leqq c<1$ the remainder after n terms of the series

$$1^rx+2^rx^2+3^rx^3+...$$

is less than $$(n+1)^rx^{n+1}/\{1-(1+1/n)^rx\}.$$

State any restrictions on n and r.

6. If $u_n=\dfrac{1}{x+2n-1}+\dfrac{1}{x+2n}-\dfrac{1}{x+n}$, $x\geqq 0$, then $\displaystyle\sum_{n=1}^{\infty}u_n=\log 2$.

7. If $u_n(x)=x^2(x^2+2^2)(x^2+4^2)\,...\,[x^2+(2n-2)^2]/(2n)!$ the series

$$1+\sum_1 u_n(x)$$

converges uniformly for every x.

8. If $u_n(x)=1/(n^3+n^4x^2)$ the series $\displaystyle\sum_1 u_n(x)$ converges uniformly for every x.

9. If Σa_n is convergent the series

$$\sum a_n\frac{x^n}{1+x^n},\quad \sum a_n\frac{x^n}{1+x^{2n}},\quad \sum a_n\frac{nx^n(1-x)}{1-x^n},\quad \sum\frac{2na_nx^n(1-x)}{1-x^{2n}}$$

converge uniformly for $0\leqq x\leqq 1$. (Hardy.)

10. If the series Σa_n is convergent and if c is neither zero nor a negative integer show that the series

$$\sum_{n=1}^{\infty}a_n\frac{c(c+1)\,...\,(c+n-1)}{x(x+1)\,...\,(x+n-1)}$$

is uniformly convergent if $x-c\geqq\delta>0$ and x neither zero nor a negative integer. (If c and x are complex, then the real part of $x-c$ will be greater than δ.)

11. If $\Sigma u_n(x)$ converges uniformly for $a<x<b$ and if each of the functions $u_n(x)$ tends to a_n when x tends to a from within the interval (a, b), show (i) that the series Σa_n converges, say to A, and (ii) that

$$\mathop{\mathcal{L}}_{x\to a} \Sigma u_n(x) = A = \Sigma \mathop{\mathcal{L}}_{x\to a} u_n(x).$$

(Note that the *value* of $u_n(x)$ for $x=a$ is not in question; $u_n(x)$ may or may not be defined for $x=a$ so long as $u_n(x)$ tends to a limit when $x\to a$.)

[With the usual notation, m can be chosen because of the uniform convergence so that, for every x such that $a<x<b$, and for $n \geqq m$, $|{}_pR_n(x)|<\epsilon$, and therefore, if n and p be kept constant while $x\to a$,

$$\mathop{\mathcal{L}}_{x\to a} |{}_pR_n(x)| \leqq \epsilon, \text{ that is, } |a_{n+1}+a_{n+2}+\ldots+a_{n+p}| \leqq \epsilon.$$

Thus Σa_n is convergent. Now take the complete remainders $R_m(x)$ and R_m in the series $\Sigma u_n(x)$ and Σa_n respectively, and write

$$\sum_1^\infty u_n(x) - \sum_1^\infty a_n = \sum_1^m \{u_n(x)-a_n\} + R_m(x) - R_m;$$

we can choose m so that both $|R_m(x)|$ and $|R_m|$ are less than ϵ $(a<x<b)$ and then, m being kept fixed, h may be chosen so that

$$\left|\sum_1^m \{u_n(x)-a_n\}\right| < \epsilon \text{ if } x-a<h,$$

and therefore $$\left|\sum_1^\infty u_n(x) - \sum_1^\infty a_n\right| < 3\epsilon \text{ if } x-a<h,$$

that is, $\Sigma u_n(x)\to\Sigma a_n$ if $x\to a$.

Of course a like theorem holds for $x\to b$.]

12. $\mathop{\mathcal{L}}_{x\to 1} (x-x^4+x^9-x^{16}+\ldots)=\frac{1}{2}$.

[If $f(x)=x-x^4+\ldots$ the coefficient a_n of x^n in $f(x)(1-x)^{-1}$ is 1 or 0 according as the greatest integer in $\sqrt{n}$ is odd or even. If $[\sqrt{n}]=\lambda=$ greatest integer in $\sqrt{n}$, then

$$a_1+a_2+\ldots+a_n=\tfrac{1}{2}\{1-(-1)^\lambda\}(n+1)+\tfrac{1}{2}(-1)^\lambda\lambda(\lambda+1).$$

$$\mathop{\mathcal{L}}_{x\to 1}\frac{f(x)(1-x)^{-2}}{(1-x)^{-2}} = \mathop{\mathcal{L}}_{n\to\infty}\frac{a_1+a_2+\ldots+a_n}{n+1}=\tfrac{1}{2}.]$$

13. Lambert's Series is $\dfrac{x}{1-x}+\dfrac{x^2}{1-x^2}+\dfrac{x^3}{1-x^3}+\ldots$; prove that it converges if $|x|<1$. Express it as a double series and show that it may be transformed into the series (Clausen's Series)

$$x\frac{1+x}{1-x}+x^4\frac{1+x^2}{1-x^2}+x^9\frac{1+x^3}{1-x^3}+\ldots.]$$

14. Prove that Lambert's Series may be expressed in the form

$$x\theta(1)+x^2\theta(2)+x^3\theta(3)+\ldots+x^n\theta(n)+\ldots$$

where $\theta(n)$ is the number of divisors of n, including 1 and n.

15. If $|x|<1$, show that

$$\frac{x}{1+x^2}+\frac{x^2}{1+x^4}+\frac{x^3}{1+x^6}+\ldots=\frac{x}{1-x}-\frac{x^3}{1-x^3}+\frac{x^5}{1-x^5}-\ldots.$$

16. If $|x|<1$, show that

$$\frac{x}{1+x^2}+\frac{x^3}{1+x^6}+\frac{x^5}{1+x^{10}}+\ldots=\frac{x}{1-x^2}-\frac{x^3}{1-x^6}+\frac{x^5}{1-x^{10}}-\ldots.$$

17. If $|x|<1$, show that

$$\frac{x}{1+x}-\frac{2x^2}{1+x^2}+\frac{3x^3}{1+x^3}-\ldots=\frac{x}{(1+x)^2}-\frac{x^2}{(1+x^2)^2}+\frac{x^3}{(1+x^3)^2}-\ldots.$$

18. If $f(x)=\sum_{n=1}^{\infty}a_nx^n$ and $g(x)=\sum_{n=1}^{\infty}b_nx^n$, both series converging for $|x|<1$, show that

$$\sum_{n=1}^{\infty}b_nf(x^n)=\sum_{n=1}^{\infty}a_ng(x^n). \qquad \text{(Knopp.)}$$

67. Series of Complex Terms. If b and c are real numbers and i is "the imaginary unit $\sqrt{(-1)}$," the number a where $a=b+ci$ is called a complex number. The student will be supposed to be familiar with the usual nomenclature and with the method of representing complex numbers on the Argand Diagram, as well as with the laws of operation.

For definiteness, it may be noted that when b and c are given and the numbers r and θ are determined by the equations

$$r\cos\theta=b,\ r\sin\theta=c$$

subject to the conditions that r is *positive* and $-\pi<\theta\leqq\pi$, the number r is called the *modulus* of a or $b+ic$ and is denoted by $|a|$ or $|b+ic|$, being equal to the positive value of $(b^2+c^2)^{\frac{1}{2}}$, while θ is called *the amplitude* of a. If θ is one solution of the equations so is $\theta+2n\pi$ where n is any positive or negative integer; the value of θ as above restricted is called the *principal value* of the amplitude of a. Again, the principal value is often taken to satisfy the condition $0\leqq\theta<2\pi$ but, unless otherwise specified, the principal value will be supposed to be such that $-\pi<\theta\leqq\pi$.

If b_n and c_n are real numbers and $a_n=b_n+ic_n$ the series

$$\Sigma a_n \text{ or } \Sigma(b_n+ic_n)$$

is called a series of complex terms or a complex series. It is plain that if the series Σb_n and Σc_n are both convergent the

number $\Sigma(b_n + ic_n)$ or Σa_n is a definite number and in this case the series Σa_n is said to converge. If either Σb_n or Σc_n is not convergent the number $\Sigma(b_n + ic_n)$ is not a definite number and Σa_n is said to be not convergent. The convergence of Σa_n may therefore be tested by seeing whether Σb_n and Σc_n are *both* convergent.

The most important type is the *absolutely* convergent series. The series Σa_n or $\Sigma(b_n + ic_n)$ is said to be absolutely convergent if each of the series Σb_n and Σc_n is absolutely convergent, and it is easily proved that Σa_n is absolutely convergent if, and only if, $\Sigma | a_n |$ is convergent. For

$$| a_n | = (b_n^2 + c_n^2)^{\frac{1}{2}} \leqq | b_n | + | c_n |$$

and therefore $\Sigma | a_n |$ converges if both $\Sigma | b_n |$ and $\Sigma | c_n |$ converge.

Again, $\quad | b_n | \leqq (b_n^2 + c_n^2)^{\frac{1}{2}} = | a_n |, \quad | c_n | \leqq | a_n |$

so that both $\Sigma | b_n |$ and $\Sigma | c_n |$ converge if $\Sigma | a_n |$ converges.

The convergence of $\Sigma | a_n |$ may be tested by the rules for series of positive terms.

For example, the series Σx^n, where x is complex, converges absolutely if $| x | < 1$ because $| a_{n+1} | \div | a_n | = | x |$.

The series $\Sigma x^n/n!$ converges absolutely for every x; because

$$\frac{| a_{n+1} |}{| a_n |} = \left| \frac{a_{n+1}}{a_n} \right| = \frac{| x |}{n+1} \to 0 \text{ if } n \to \infty .$$

When a_n/a_{n+1} can be expressed in the form

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\alpha + i\beta}{n} + \frac{A_n + iB_n}{n^2}$$

where $| A_n |$ and $| B_n |$ are bounded, we have

$$\left| \frac{a_n}{a_{n+1}} \right| = \left\{ \left(1 + \frac{\alpha}{n} + \frac{A_n}{n^2}\right)^2 + \left(\frac{\beta}{n} + \frac{B_n}{n^2}\right)^2 \right\}^{\frac{1}{2}}$$

$$= 1 + \frac{\alpha}{n} + \frac{C_n}{n^2}, \quad |C_n| \text{ bounded,}$$

and therefore (§ 60) Σa_n will converge absolutely if $\alpha > 1$ but $\Sigma |a_n|$ will diverge if $\alpha \leqq 1$. Instead of $(A_n + iB_n)/n^2$ we might have $(A_n + iB_n)/n^\lambda$, $\lambda > 1$.

Ex. If α, β, γ, x are complex, $\alpha = \alpha_1 + i\alpha_2$, $\beta = \beta_1 + i\beta_2$, $\gamma = \gamma_1 + i\gamma_2$, the hypergeometric series will converge absolutely when $x = 1$ if $(\gamma_1 - \alpha_1 - \beta_1)$, that is, if the Real Part of $(\gamma - \alpha - \beta)$, is positive.

Tannery's Theorem, § 63, holds for complex terms as for real; the proof needs no change when the terms of $\Sigma u_r(n)$ are complex.

Derangement of Series. Here too no change is required; when the conditions (i) and (ii) of § 66 are satisfied the proof is the same as when the terms are real.

68. The Exponential Function. Let $F(n) = (1 + z/n)^n$ where n is a positive integer and z is complex, $z = x + iy$, x and y being real; it will be proved that, when n tends to infinity, $F(n)$ tends to the limit expressed by the series

$$1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \frac{z^r}{r!} + \dots = \varphi(z), \text{ say}, \quad \dots\dots\dots\dots(1)$$

which converges absolutely for every z.

Expand $(1 + z/n)^n$ by the binomial theorem, which is applicable when n is a positive integer, and express the coefficients as in § 48 of the *Elementary Treatise*; thus

$$F(n) = 1 + z + \left(1 - \frac{1}{n}\right)\frac{z^2}{2!} + \dots + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{r-1}{n}\right)\frac{z^r}{r!} + \dots$$
$$+ \dots + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{n-1}{n}\right)\frac{z^n}{n!}. \quad (2)$$

Let $u_r(n)$ be the term of the expansion that contains z^r, and let $|z| = a$; then we have

(i) $\underset{n\to\infty}{\mathcal{L}}\, u_r(n) = z^r/r!$, when r is fixed;

(ii) $|u_r(n)| \leqq a^r/r!$; (iii) $\Sigma a^r/r!$ is convergent.

The conditions of Tannery's Theorem, § 63, are therefore satisfied and

$$\underset{n\to\infty}{\mathcal{L}}\, F(n) = \sum_{r=0}^{\infty} \frac{z^r}{r!} = \varphi(z) \quad \dots\dots\dots\dots\dots\dots(3)$$

and this series converges *absolutely* for every z.

The limit may, however, be expressed in another form, namely

$$\underset{n\to\infty}{\mathcal{L}}\, F(n) = e^x(\cos y + i \sin y).$$

For, if $1 + x/n = r\cos\theta$ and $y/n = r\sin\theta$ (n a positive integer) so that $1 + z/n = (1 + x/n) + iy/n = r(\cos\theta + i\sin\theta)$, $r\cos\theta$ will, for large values of n, be positive while $r\sin\theta$ will

have the same algebraic sign as y; hence θ may be chosen so that it lies between $-\pi/2$ and $\pi/2$. By De Moivre's Theorem

$$(1+z/n)^n = r^n(\cos\theta + i\sin\theta)^n = r^n(\cos n\theta + i\sin n\theta).$$

Now,

$$r^n = \left\{\left(1+\frac{x}{n}\right)^2 + \frac{y^2}{n^2}\right\}^{\frac{n}{2}} = \left(1+\frac{x}{n}\right)^n \left\{1+\frac{y^2}{(n+x)^2}\right\}^{\frac{n}{2}}.$$

But $ny^2/(n+x)^2 \to 0$ when $n\to\infty$ and therefore by § 25, Ex. 2,

$$\{1+y^2/(n+x)^2\}^{\frac{n}{2}} \to 1 \text{ when } n\to\infty,$$

so that $$r^n \to e^x \text{ when } n\to\infty.$$

Again, $\tan\theta = y/(n+x)$ and therefore $\theta\to 0$ when $n\to\infty$. Also,

$$n\theta = \frac{\theta}{\tan\theta}\cdot\frac{ny}{n+x} = \frac{\theta}{\tan\theta}\cdot\frac{y}{1+x/n} \to y \text{ when } n\to\infty$$

if we assume the usual theorem that $\theta/\tan\theta \to 1$ when $\theta\to 0$.

Hence $$\mathop{\mathcal{L}}_{n\to\infty} F(n) = \mathop{\mathcal{L}}_{n\to\infty}\left(1+\frac{x+iy}{n}\right)^n = e^x(\cos y + i\sin y) \quad \text{.......(4)}$$

and therefore

$$\mathcal{L}\left(1+\frac{z}{n}\right)^n = e^x(\cos y + i\sin y) = \sum_{r=0}^{\infty}\frac{z^r}{r!} = \varphi(z).$$

When z is real, $z=x$, the series is equal to e^x; the definition of the exponential function is now extended to complex values by saying that e^z means the series $\Sigma z^r/r!$ or $\varphi(z)$ or the function $e^x(\cos y + i\sin y)$ when $z = x+iy$.

69. Trigonometric and Hyperbolic Functions. If z_1 and z_2 are any two complex numbers the power e^z satisfies the index law

$$e^{z_1}\times e^{z_2} = e^{z_1+z_2},$$

as may be verified either by finding the product $\varphi(z_1)\times\varphi(z_2)$, which is easily seen to be $\varphi(z_1+z_2)$, or by taking the product of e^{z_1} and e^{z_2} when these are expressed in the form

$$e^{x_1}(\cos y_1 + i\sin y_1) \text{ and } e^{x_2}(\cos y_2 + i\sin y_2).$$

If n is any positive or negative integer $(e^z)^n = e^{nz}$, but when n is not integral or not real the function is no longer single-valued. (See § 72.)

Suppose $x=0$; the series for $\varphi(z)$ then gives, by equating real and imaginary parts,

$$\cos y = 1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \frac{y^6}{6!} + \dots$$

$$\sin y = y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \dots$$

which are the usual power series for $\cos y$ and $\sin y$; by equation (4), § 68, y is the number of radians of angle.

In the equation $e^{iy} = \cos y + i \sin y$ put $-y$ for y; the equations $e^{iy} = \cos y + i \sin y$, $e^{-iy} = \cos y - i \sin y$

give $$\cos y = \frac{e^{iy} + e^{-iy}}{2}, \quad \sin y = \frac{e^{iy} - e^{-iy}}{2i} \quad \dots\dots\dots\dots\dots(1)$$

which are known as *Euler's expressions for* $\cos y$ *and* $\sin y$.

The direct trigonometric functions of a complex number z which, as yet, have no existence, are now introduced by definitions suggested by equations (1), namely,

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i},$$

while the other functions $\tan z$, $\cot z$, $\operatorname{cosec} z$, $\sec z$ are defined by the equations $\tan z = \sin z / \cos z$, ... $\sec z = 1/\cos z$, which hold for real angles.

Again, $\sin^2 z + \cos^2 z = \left(\frac{e^{iz} - e^{-iz}}{2i}\right)^2 + \left(\frac{e^{iz} + e^{-iz}}{2}\right)^2 = 1,$

the fundamental identity for real angles. Similarly it is seen that $$1 + \tan^2 z = \sec^2 z, \quad 1 + \cot^2 z = \operatorname{cosec}^2 z,$$

and it is also verified very easily that the Addition Theorems for $\cos(z_1 \pm z_2)$ and $\sin(z_1 \pm z_2)$ hold also for complex values.

It should be noted, however, that, when z is complex, the familiar relations $|\sin z| \leqq 1$, $|\cos z| \leqq 1$ are no longer true in general.

Periodicity of e^z. The function e^z is periodic, with the pure imaginary period $2\pi i$; for if n is any positive or negative integer,

$$e^{z+n \cdot 2\pi i} = e^z \times e^{2n\pi i} = e^z(\cos 2n\pi + i \sin 2n\pi) = e^z.$$

The trigonometric functions have, however, the same *real* periods as when the angles are real; for

$$e^{\pm i(z+n \cdot 2\pi)} = e^{\pm iz} \cdot e^{\pm 2n\pi i} = e^{\pm iz},$$

so that an increase or decrease of the argument z in $\sin z$ or $\cos z$ by 2π makes no change in the value of $\sin z$ or $\cos z$.

The zeros of the trigonometric functions are the same as when the angles are real. For example, if $z = x + iy$ and $\sin z = 0$ we must have

$$e^{ix-y} = e^{-ix+y} \quad \text{or} \quad e^{2y} = e^{2ix} = \cos 2x + i \sin 2x.$$

This equation requires, since x and y are real, $e^{2y} = 1$ or $y = 0$ and $\cos 2x = 1$, $\sin 2x = 0$ so that $x = n\pi$, $n = 0, \pm 1, \pm 2, \ldots$

The Hyperbolic Functions. The definitions are the same as when z is real (*E.T.* p. 140); for example,

$$\cosh z = \tfrac{1}{2}(e^z + e^{-z}), \quad \sinh z = \tfrac{1}{2}(e^z - e^{-z}).$$

The relations

$$\cos iz = \cosh z, \quad \sin iz = i \sinh z, \quad \sinh iz = i \sin z$$

should be noted as they are often useful.

Ex. 1. Show that, x and y being real,

(i) $\cos(x \pm iy) = \cos x \cosh y \mp i \sin x \sinh y$;

(ii) $\sin(x \pm iy) = \sin x \cosh y \pm i \cos x \sinh y$;

(iii) $\tan(x \pm iy) = \dfrac{\sin 2x \pm i \sinh 2y}{\cos 2x + \cosh 2y}$.

Ex. 2. If x and y are real prove that

(i) $|\sin(x \pm iy| = (\sin^2 x + \sinh^2 y)^{\frac{1}{2}} \leqq \cosh y$ but $\geqq \sinh |y|$.

(ii) $|\cos(x \pm iy| = (\cos^2 x + \sinh^2 y)^{\frac{1}{2}} \leqq \cosh y$ but $\geqq \sinh |y|$.

Ex. 3. If $x + iy = \tan(u + iv)$ where x, y, u, v are all real, show that

(i) $x^2 + y^2 = (\cosh 2v - \cos 2u)/(\cosh 2v + \cos 2u)$;

(ii) $\tan 2u = \dfrac{2x}{1 - x^2 - y^2}$; (iii) $\tanh 2v = \dfrac{2y}{1 + x^2 + y^2}$;

(iv) $e^{4v} = \{x^2 + (y+1)^2\}/\{x^2 + (y-1)^2\}$.

Ex. 4. If z is complex, $|e^z - 1| < |z|(1 + \frac{1}{2}|z|e^{|z|})$.

Let $|z| = r$; then $|e^z - 1|$ is less than or equal to

$$r + \tfrac{1}{2}r^2\left\{1 + \frac{r}{3} + \frac{r^2}{3\,.\,4} + \ldots + \frac{r^n}{3\,.\,4 \ldots (n+2)} + \ldots\right\}$$

$$< r + \tfrac{1}{2}r^2\left\{1 + r + \frac{r^2}{2!} + \ldots + \frac{r^n}{n!} + \ldots\right\} = r\left[1 + \tfrac{1}{2}re^r\right].$$

70. Logarithms. If $e^w = z$, where w and z are complex, w is defined to be *a logarithm of* z; when z is real there is only one real logarithm of z, but when z is complex new considerations come into play.

Let $z=x+iy=r(\cos\theta+i\sin\theta)$, where $-\pi<\theta\leqq\pi$, and let $w=u+iv$, where u and v like x and y are real; then

$$e^w=e^{u+iv}=x+iy=r(\cos\theta+i\sin\theta),$$

or
$$e^u(\cos v+i\sin v)=r(\cos\theta+i\sin\theta),$$

and therefore $\quad e^u=r,\ v=\theta+2n\pi,\ n=0,\ \pm1,\ \pm2,\ \dots$.

Hence, since r is positive, $u=\log r$, a real number, and

$$w=u+iv=\log r+i(\theta+2n\pi). \qquad \text{.....................(1)}$$

This value of w is "the general logarithm of z" and is denoted by $\operatorname{Log} z$, so that $\operatorname{Log} z$ has an unlimited number of values that differ by multiples of $2\pi i$. The value of w for which n is zero is called the principal value of the logarithm of z and is denoted by $\log z$. Hence

$$\operatorname{Log} z=\log z+2n\pi i,\ n=0,\ \pm1,\ \pm2,\ \dots \qquad \text{.........(2)}$$

If $z=x+iy$, then $r=(x^2+y^2)^{\frac{1}{2}}$, $\log r=\frac{1}{2}\log(x^2+y^2)$ while θ, the amplitude of z, is the angle which satisfies the equations $r\cos\theta=x$, $r\sin\theta=y$ *and also* the inequalities $-\pi<\theta\leqq\pi$.

If $\theta_1,\theta_2,\dots,\theta_m$ are the principal values of the amplitudes of $z_1,z_2,\dots,z_m$ and φ the principal value of the amplitude of the product $z_1,z_2,\dots,z_m$ then φ is not, in general, equal to $\theta_1+\theta_2+\dots+\theta_m$ but to $\theta_1+\theta_2+\dots+\theta_m+2k\pi$, where k may be 0 but, in general, is a positive or negative integer which must be chosen so that $-\pi<\varphi\leqq\pi$. Hence

$$\log(z_1z_2\dots z_m)=\log z_1+\log z_2+\dots+\log z_m+2k\pi i.$$

Ex. Let $m=2$.

$$\theta_1=\frac{\pi}{6},\ \theta_2=\frac{2\pi}{3},\ k=0;\ \theta_1=\theta_2=\frac{2\pi}{3},\ k=-1,\ \varphi=-\frac{2\pi}{3};$$

$$\theta_1=-\frac{\pi}{2},\ \theta_2=-\frac{2\pi}{3},\ k=1,\ \varphi=\frac{5\pi}{6}.$$

71. Inverse Trigonometric Functions. If x is real and $\tan y=x$ we find by expressing $\tan y$ in terms of e^{iy} that

$$\tan y=\frac{e^{iy}-e^{-iy}}{i(e^{iy}+e^{-iy})}=x,\quad e^{2iy}=\frac{1+ix}{1-ix},$$

and therefore
$$y=\frac{1}{2i}\operatorname{Log}\left(\frac{1+ix}{1-ix}\right). \qquad \text{.......................(1)}$$

Let $(1+ix)/(1-ix)=r(\cos\theta+i\sin\theta)$, $-\pi<\theta\leqq\pi$; then

$$r=1,\ \cos\theta=\frac{1-x^2}{1+x^2},\ \sin\theta=\frac{2x}{1+x^2},\ \tan\tfrac{1}{2}\theta=x$$

and
$$\frac{1}{2i}\operatorname{Log}\left(\frac{1+ix}{1-ix}\right)=\frac{1}{2i}\times i(\theta+2n\pi)=\tfrac{1}{2}\theta+n\pi.$$

The principal value of y is therefore $\frac{1}{2}\theta$ and is the value to which the symbol $\tan^{-1}x$ is restricted (*E.T.* p. 133).

The general value of $\tan^{-1}z$, when z is complex, is defined to be

$$\frac{1}{2i}\operatorname{Log}\left(\frac{1+iz}{1-iz}\right). \quad \text{.......................(2)}$$

Thus if $x+iy=\tan(u+iv)$ so that $u+iv$ is a value of $\tan^{-1}(x+iy)$, it follows from § 69, Ex. 3, that

$$u+iv=n\pi+\tfrac{1}{2}\tan^{-1}\left(\frac{2x}{1-x^2-y^2}\right)+i\,.\,\tfrac{1}{4}\log\left\{\frac{x^2+(y+1)^2}{x^2+(y-1)^2}\right\} \quad (3)$$

where $n=0,\ \pm1,\ \pm2,\dots$, and therefore, as when z is real, the values of $\tan^{-1}z$ differ by multiples of π.

As we shall make very little use of these inverse functions it is sufficient to state that all can be expressed in the form $\alpha+i\beta$ where α and β are real functions of the real variables x and y. For fuller information the student may consult Chrystal's *Algebra*, Chapter XXIX, or Hobson's *Trigonometry*, Chapter XVI.

Ex. If $\sin^{-1}(x+iy)=\alpha+i\beta$, $x+iy=\sin\alpha\cosh\beta+i\cos\alpha\sinh\beta$, prove that

(i) $\cosh\beta=\frac{1}{2}\{(x+1)^2+y^2\}^{\frac{1}{2}}+\frac{1}{2}\{(x-1)^2+y^2\}^{\frac{1}{2}}=u$;

(ii) $\sin\alpha=\frac{1}{2}\{(x+1)^2+y^2\}^{\frac{1}{2}}-\frac{1}{2}\{(x-1)^2+y^2\}^{\frac{1}{2}}=v$;

(iii) $\sin^{-1}(x+iy)=n\pi+(-1)^n\sin^{-1}v+i\,.\,(-1)^n\log\{u+(u^2-1)^{\frac{1}{2}}\}$.

72. The Generalised Power. The power z^n is defined for all values of z and n, real or complex, by the equation

$$z^n=e^{n\operatorname{Log}z}$$

and $e^{n\log z}$ is the principal value of z^n.

The general power z^n is single-valued if, and only if, n is zero or an integer, positive or negative; because

$$n\operatorname{Log}z=n\log z+n\,.\,2k\pi i,\quad k=0,\ \pm1,\pm2,\dots,$$

and in this case $e^{2nk\pi i}$ is unity. If n is a rational fraction, $n=\pm p/q$, where p and q are positive integers and p/q is in its

lowest terms, z^n has q different values. In all other cases the general power z^n has an unlimited number of different values.

Ex. 1. (i) $\log(-1)=\pi i$; (ii) $\log i=\frac{\pi}{2}i$; (iii) $\log(-i)=-\frac{\pi}{2}i$.

Ex. 2. The principal value of i^i is $e^{-\pi/2}$.

Ex. 3. If $\tan x=c\sin\alpha/(1-c\cos\alpha)$, show that if $x=0$ when $c=0$

$$x=\frac{1}{2i}\log\left(\frac{1-ce^{-i\alpha}}{1-ce^{i\alpha}}\right).$$

Ex. 4. If x and k are complex, $|x|<1$ and $k=\alpha+i\beta$ where α and β are real, find the modulus of the principal value of $(1+x)^k$.

The principal value of $\log(1+x)$ is that value which is zero when x is zero; for if $x=r(\cos\theta+i\sin\theta)$, $-\pi<\theta\leqq\pi$, and

$$1+r\cos\theta=\rho\cos\varphi,\quad r\sin\theta=\rho\sin\varphi,$$

$\cos\varphi$ is positive and therefore $-\pi/2<\varphi<\pi/2$ and $\varphi=0$ when $r=0$.

Next, $$\log(1+x)=\log\varrho+i\varphi,$$

so that $$k\log(1+x)=(\alpha\log\varrho-\beta\varphi)+i(\beta\log\varrho+a\varphi)$$

and therefore $$|(1+x)^k|=e^{a\log\rho-\beta\phi}=\varrho^a e^{-\beta\phi}$$

where $\varrho=(1+2r\cos\theta+r^2)^{\frac{1}{2}}$.

Since $$|\varphi|\leqq\pi/2,\quad|(1+x)^k|\leqq\varrho^a e^{|\beta|\frac{\pi}{2}}.$$

73. Complex Functions of a Real Variable. If u and v are real functions of the real variable x—that is, functions in which the constants are real numbers—the function $u+iv$ is called a *complex function of the real variable x.*

A polynomial $f(x)$ of degree n (n a positive integer) in which the coefficients are complex numbers may, by separating the purely real and the purely imaginary parts, be expressed in the form $u+iv$. The quotient of two such polynomials is of the form $(u_1+iv_1)/(u_2+iv_2)$ and

$$\frac{u_1+iv_1}{u_2+iv_2}=\frac{u_1u_2+v_1v_2}{u_2^2+v_2^2}+i\frac{u_2v_1-u_1v_2}{u_2^2+v_2^2}=u+iv.$$

Again, if $a=b+ic$ (b, c real) e^{ax} is of the form $u+iv$ where

$$u=e^{bx}\cos cx \text{ and } v=e^{bx}\sin cx.$$

From the definitions of logarithms and of the circular functions, direct and inverse, it will be seen that they are all of the form $u+iv$.

The definition therefore includes all the ordinary functions.

The derivative and the integral of a complex function $f(x)$ of the real variable x are defined, when $f(x)$ has been expressed

in the form $u+iv$ where u and v are real functions of x, by the equations

$$\frac{df(x)}{dx}=\frac{du}{dx}+i\frac{dv}{dx} \text{ and } \int f(x)dx=\int udx+i\int vdx+C,$$

where C is a constant, in general complex, $C=C_1+iC_2$.

The definite integral is defined by the equation

$$\int_a^b f(x)dx=\int_a^b udx+i\int_a^b vdx,$$

where the limits a and b are, of course, like x, real numbers.

The rules for differentiating a sum, a product or quotient, and a function of a function (both variables being real) will obviously be the same for the complex as for the real functions; the rule for differentiating an inverse function will be proved for the standard formulae, as the general proof really requires the concept of the complex variable.

(I) e^w, where $w=u+iv$ and u, v are real functions of the real variable x.

By definition, $e^w=e^u(\cos v+i\sin v)$; therefore

$$\frac{d\,.\,e^w}{dx}=e^u\frac{du}{dx}(\cos v+i\sin v)+e^u(-\sin v+i\cos v)\frac{dv}{dx}$$

$$=e^u(\cos v+i\sin v)\frac{du}{dx}+ie^u(\cos v+i\sin v)\frac{dv}{dx}$$

$$=e^u(\cos v+i\sin v)\left(\frac{du}{dx}+i\frac{dv}{dx}\right)$$

so that $\dfrac{d\,.\,e^w}{dx}=e^w\dfrac{dw}{dx}$.

We thus have the same rule as if w were real.

(II) Trigonometric Functions. By expressing $\sin w$, $\cos w$, $\tan w$, etc., in terms of e^{iw} and applying (I) it is readily found that the derivatives have the same form as when w is real.

For example,

$$\frac{d\,.\sin w}{dx}=\frac{1}{2i}\frac{d(e^{iw}-e^{-iw})}{dx}=\tfrac{1}{2}(e^{iw}+e^{-iw})\frac{dw}{dx},$$

so that $\dfrac{d\,.\sin w}{dx}=\cos w\dfrac{dw}{dx}$.

(III) $\log w$. In this case let $w = \varrho(\cos\varphi + i\sin\varphi)$ where ϱ and φ are real functions of x; then $\log w = \log\varrho + i\varphi$ so that

$$\frac{d \,.\log w}{dx} = \frac{1}{\varrho}\frac{d\varrho}{dx} + i\frac{d\varphi}{dx}.$$

Now
$$\frac{dw}{dx} = \frac{d\varrho}{dx}(\cos\varphi + i\sin\varphi) + i\varrho(\cos\varphi + i\sin\varphi)\frac{d\varphi}{dx}$$

$$= w\left(\frac{1}{\varrho}\frac{d\varrho}{dx} + i\frac{d\varphi}{dx}\right)$$

and therefore

$$\frac{d \,.\log w}{dx} = \frac{1}{w}\frac{dw}{dx}.$$

Since $\operatorname{Log} w = \log w + 2n\pi i$ the derivative of $\operatorname{Log} w$ is the same as that of $\log w$.

(IV) Inverse Trigonometric functions. The forms are the same as when w is real; for example

$$\frac{d \,.\tan^{-1} w}{dx} = \frac{1}{2i}\frac{d}{dx}\log\left(\frac{1+iw}{1-iw}\right) = \frac{1}{1+w^2}\frac{dw}{dx}.$$

(V) w^n. Let $\log w$ be the principal value of $\operatorname{Log} w$ and $w^n = e^{n\log w}$ where w is a complex function of the real variable x and n a complex constant. We find by (I) and (III)

$$\frac{d \,. w^n}{dx} = e^{n\log w}\frac{d(n\log w)}{dx} = w^n \,.\, \frac{n}{w}\frac{dw}{dx},$$

so that
$$\frac{d \,. w^n}{dx} = nw^{n-1}\frac{dw}{dx}.$$

Cor. If the same value of $\operatorname{Log} w$ is used for w^n and w^{n-1}

$$\frac{d \,. w^n}{dx} = nw^{n-1}\frac{dw}{dx}.$$

(VI) In respect of integration we may assume (as will be fairly evident from consideration of the definite integral as the limit of a sum) that

$$\left|\int_a^b (u+iv)dx\right| \leqq \int_a^b |u+iv|\,dx$$

or
$$\left|\int_a^b w\,dx\right| \leqq \int_a^b |w|\,dx,$$

where a and b are real numbers and $a < b$.

It is also assumed that the integral of $df(x)/dx$ is $f(x) +$ constant.

Ex. Find the integrals of $e^{ax}\cos bx$ and $e^{ax}\sin bx$, (a, b real).

$$\int e^{(a+ib)x}\,dx = \frac{e^{(a+ib)x}}{a+ib} = \frac{e^{ax}(\cos bx + i\sin bx)}{a+ib}$$

and
$$\int e^{(a+ib)x}\,dx = \int e^{ax}\cos bx\,dx + i\int e^{ax}\sin bx\,dx.$$

Equate real and imaginary parts.

74. Logarithmic and Binomial Series. The series for $\log(1+x)$ and $(1+x)^k$, where x and k are complex, will now be established.

I. $\log(1+x)$. When t is real, $0 \leqq t \leqq 1$, and x complex, $|x| < 1$, the principal value of $\log(1+xt)$ is that for which x (or t) is zero (§ 72, Ex. 4).

Now, let $|x| = \varrho < 1$ and let t be real; the binomial $1+xt$ cannot vanish if $0 \leqq t \leqq 1$, and therefore

$$\int_0^1 \frac{x\,dt}{1+xt} = \text{Principal value of Log}\,(1+x).$$

Again, by elementary algebra, we have

$$\frac{x}{1+xt} = \sum_{r=1}^{n}(-1)^{r-1}x^r t^{r-1} + (-1)^n \frac{x^{n+1}t^n}{1+xt},$$

and therefore, $\log(1+x)$ denoting the principal value of Log $(1+x)$,

$$\log(1+x) = \sum_{r=1}^{n}(-1)^{r-1}\frac{x^r}{r} + (-1)^n R_n(x),\ |x| = \varrho < 1, \quad \text{......(1)}$$

where
$$R_n(x) = \int_0^1 \frac{x^{n+1}t^n}{1+xt}\,dt.$$

Now $|1+xt| \geqq 1-\varrho > 0$ when $|x| = \varrho < 1$ and $0 \leqq t \leqq 1$; hence

$$|R_n(x)| < \frac{\varrho^{n+1}}{1-\varrho}\int_0^1 t^n dt, \quad |R_n(x)| < \frac{\varrho^{n+1}}{1-\varrho}\cdot\frac{1}{n+1} \quad \text{.......(2)}$$

and therefore $R_n(x) \to 0$ when $n \to \infty$. We thus find that

$$\log(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \dots,\ |x| < 1, \quad \text{.......(3)}$$

so that the series for the principal value of $\log(1+x)$ when x is complex and $|x| < 1$ is the same as that for $\log(1+x)$ when x is real and $|x| < 1$.

Ex. 1. If $|x| < \frac{1}{2}$, $|\log(1+x)| \leqq 2|x|$.

In the inequality (2) let

$$n=1;\ \text{then}\ 1-\varrho > \tfrac{1}{2}\ \text{and}\ |R_1(x)| \leqq |x|^2 \leqq |x|,$$

so that
$$|\log(1+x)| \leqq |x| + |R_1(x)| \leqq 2|x|.$$

(The sign $=$ occurs only for $x=0$.)

Ex. 2. $|\log(1+x)| \leqq |x| + \frac{1}{2}|x|^2 + \frac{1}{3}|x|^3 + \dots$ and therefore $|\log(1+x)| \leqq -\log(1-|x|)$, $|x|<1$.

Ex. 3. Show that $|-\frac{1}{2}+\frac{1}{3}x-\frac{1}{4}x^2+\frac{1}{5}x^3-\dots|<1$ if $|x|<\frac{1}{2}$ and deduce that $\log(1+x)=x+\theta x^2$ where $|\theta|<1$ if $|x|<\frac{1}{2}$.

II. $(1+x)^k$. Suppose that x and k are complex, $|x|<1$, and let t be a real variable, $0 \leqq t \leqq 1$; the principal value of $(1+xt)^k$ is that value which is equal to 1 when $t=0$ and the principal value of $(1+x)^k$ is that value which is equal to 1 when $x=0$.

Let $F(t)$ be a complex function of the real variable t defined for the range $0 \leqq t \leqq 1$ by the equation

$$F(t) = \sum_{r=0}^{n-1} \binom{k}{r} x^r (1+xt)^{k-r}(1-t)^r \quad \dots\dots (4)$$

where

$$\binom{k}{r} = \frac{k(k-1)(k-2)\dots(k-r+1)}{1.2.3\dots.r}; \quad \binom{k}{0} = 1. \quad \dots(5)$$

Now the rules of differentiation with respect to the real variable t are the same as if x and k were real; differentiating $F(t)$ we find (compare *E.T.* pp. 390, 391) that

$$\frac{dF(t)}{dt} = n\binom{k}{n} x^n (1+xt)^{k-n}(1-t)^{n-1}. \quad \dots(6)$$

Since $|x|<1$ the binomial $(1+xt)$ cannot be zero for $0 \leqq t \leqq 1$ and therefore every power of $(1+xt)$ is finite; we may therefore integrate with respect to t from 0 to 1 and then equation (6) gives

$$F(1) - F(0) = n\binom{k}{n} x^n \int_0^1 (1+xt)^{k-n}(1-t)^{n-1}\,dt = R_n(x). \dots(7)$$

Again, from (4) we find that

$$F(1) - F(0) = (1+x)^k - \sum_{r=0}^{n-1}\binom{k}{r} x^r,$$

and therefore, by (7),

$$(1+x)^k = \sum_{r=0}^{n-1}\binom{k}{r} x^r + R_n(x)$$
$$= 1 + kx + \frac{k(k-1)}{1.2}x^2 + \dots$$
$$+ \frac{k(k-1)\dots(k-n+2)}{1.2\dots(n-1)} x^{n-1} + R_n(x) \quad \dots\dots\dots\dots(8)$$

where $R_n(x)$ is given by the integral in equation (7).

It may now be shown by a method analogous to that given on p. 394 of the *Elementary Treatise* that $R_n(x) \to 0$ when $n \to \infty$ if $|x|<1$.

Let $k=\alpha+i\beta$, $x=\varrho(\cos\theta+i\sin\theta)$, $-\pi<\theta\leqq\pi$, and put the integral for $R_n(x)$ in the form

$$R_n(x)=n\binom{k}{n}x^n\int_0^1(1+xt)^{k-1}\left(\frac{1-t}{1+xt}\right)^{n-1}dt.$$

Now, since t is real and $|x|=\varrho<1$ we have, if $0\leqq t\leqq 1$,

$$|1+xt|\geqq 1-\varrho t\geqq 1-t \text{ so that } |(1-t)\div(1+xt)|\leqq 1,$$

and therefore $|\{(1-t)/(1+xt)\}^{n-1}|\leqq 1$ if $0\leqq t\leqq 1$.

Again, since $1+xt$ cannot be zero when $|x|<1$ and $0\leqq t\leqq 1$, the power $|(1+xt)^{k-1}|$ must be finite, say less than M, for $0\leqq t\leqq 1$.

Further, if $|k|=\sqrt{(\alpha^2+\beta^2|}=\kappa$, we have

$$\left|n\binom{k}{n}\right|=\kappa\left|\binom{k-1}{n-1}\right|<\kappa\,.\,\frac{(\kappa+1)(\kappa+2)\ldots(\kappa+n-1)}{1\,.\,2\ldots(n-1)}.$$

Hence $$\left|\int_0^1(1+xt)^{k-1}\left(\frac{1-t}{1+xt}\right)^{n-1}dt\right|<M\int_0^1 dt=M$$

and $$\left|R_n(x)\right|<\kappa\frac{(\kappa+1)(\kappa+2)\ldots(\kappa+n-1)}{1\,.\,2\,.\ldots(n-1)}M\varrho^n=\kappa\varrho M\,.\,a_{n-1}$$

where a_{n-1} is the nth term of the convergent series

$$1+(\kappa+1)\varrho+\frac{(\kappa+1)(\kappa+2)}{1\,.\,2}\varrho^2+\ldots.$$

Therefore, since $a_{n-1}\to 0$ when $n\to\infty$, being the nth term of a convergent series, the Remainder $R_n(x)$ in the expansion (8) also tends to zero when $n\to\infty$. The principal value of $(1+x)^k$ is therefore given by the series

$$(1+x)^k=1+kx+\frac{k(k-1)}{1\,.\,2}x^2+\ldots\quad |x|<1,$$

so that the expansion has the same form as when x and k are real.

Ex. 4. If x is complex and $f_n=\left(1+\frac{1}{n}\right)^x\left(1+\frac{x}{n}\right)^{-1}$, prove that

$$\lim_{n\to\infty}n^2(f_n-1)=\tfrac{1}{2}x(x-1).$$

If $n>|x|$, we have

$$f_n=\left(1+\frac{x}{n}+\frac{x(x-1)}{2n^2}+\ldots\right)\left(1-\frac{x}{n}+\frac{x^2}{n^2}-\ldots\right)$$

$$=1+\frac{x(x-1)}{2n^2}+\frac{A_n}{n^3},\quad |A_n|\to\tfrac{1}{6}|x(x^2-1)|,$$

so that $$\lim_{n\to\infty}n^2(f_n-1)=\tfrac{1}{2}x(x-1).$$

Cor. The series $\Sigma(f_n-1)$ converges absolutely for every value of x.

75. Uniform Convergence. When z is complex, $z=x+iy$ where x and y are real, a series $\Sigma u_n(z)$ will be of the form

$$\Sigma v_n(x, y)+i\,\Sigma w_n(x, y)$$

where $v_n(x, y)$ and $w_n(x, y)$ are real functions of the real variables x and y.

If when $|z| \leqq \varrho$ or, more generally, when z is any point within or on the boundary of a closed curve C, the term $u_n(z)$ is such that

(i) $|u_n(z)| \leqq M_n$, where M_n is a positive constant

and (ii) the series ΣM_n converges,

the series $\Sigma u_n(z)$ converges absolutely and uniformly for such values of z.

The series converges since ΣM_n converges; further, the convergence does not depend on any particular value of z so long as z is in the region specified, because the convergence of the series of constant terms ΣM_n is independent of z. The property of uniform convergence is therefore maintained.

It is not hard to state theorems corresponding to those of Dirichlet and Abel, but for these and other developments we refer to Bromwich's *Infinite Series* (2nd Ed.), Chapter X.

EXERCISES VIII.

1. Show that the points on the Argand Diagram that represent the roots of the equation $(z+1)^5=32z^5$ are concyclic.

2. If x, α, β are real, show that the two series

$$C=\sum_{r=0}^{n-1} x^r \cos(\alpha+r\beta), \quad S=\sum_{r=0}^{n-1} x^r \sin(\alpha+r\beta)$$

may be expressed in closed form by summing the geometric progression

$$\sum_{r=0}^{n-1} x^r e^{i(\alpha+r\beta)}$$

and equating real and imaginary parts.

3. $\displaystyle\sum_{n=0}^{\infty}\frac{x^n \sin(\alpha+n\beta)}{n!}=e^{x\cos\beta}\sin(\alpha+x\sin\beta).$

4. From the equation $(1-z)^{-1}=\sum_{n=0}^{\infty} z^n$, $|z|<1$, deduce that, if $|x|<1$,

(i) $\dfrac{1-x\cos\theta}{1-2x\cos\theta+x^2}=1+x\cos\theta+x^2\cos 2\theta+\ldots;$

(ii) $\dfrac{\sin\theta}{1-2x\cos\theta+x^2}=\sin\theta+x\sin 2\theta+x^2\sin 3\theta+\ldots;$

(iii) $\dfrac{1-x^2}{1-2x\cos\theta+x^2}=1+2x\cos\theta+2x^2\cos 2\theta+\ldots.$

5. Show that $(1-2x\cos\theta+x^2)^{-1}$ is equal to

$$\frac{1}{2i\sin\theta}\left(\frac{e^{i\theta}}{1-xe^{i\theta}}-\frac{e^{-i\theta}}{1-xe^{-i\theta}}\right) \text{ and deduce Ex. 4, (ii).}$$

6. If $x=r(\cos\theta+i\sin\theta)$ and $r<1$, $-\pi<\theta\leqq\pi$, prove that

$$\varrho=|1+x|=\sqrt{(1+2r\cos\theta+r^2)},$$

and $\varphi=\operatorname{amp}(1+x)$, where $\varrho\cos\varphi=1+r\cos\theta$, $\varrho\sin\varphi=r\sin\theta$,

so that $\varphi=\tan^{-1}\{r\sin\theta/(1+r\cos\theta)\}$, $-\dfrac{\pi}{2}<\varphi<\dfrac{\pi}{2}$.

Deduce from the series for $\log(1+x)$ that

(i) $\frac{1}{2}\log(1+2r\cos\theta+r^2)=\sum_{n=1}^{\infty}(-1)^{n-1}\dfrac{r^n}{n}\cos n\theta$;

(ii) $\tan^{-1}\left(\dfrac{r\sin\theta}{1+r\cos\theta}\right)=\sum_{n=1}^{\infty}(-1)^{n-1}\dfrac{r^n}{n}\sin n\theta$.

7. In Ex. 6, let $r\to 1$ and show, by Abel's Theorem, that

(i) $\log(2\cos\frac{1}{2}\theta)=\sum_{n=1}^{\infty}(-1)^{n-1}\dfrac{\cos n\theta}{n}$, $-\pi<\theta<\pi$;

or, 2θ being put in place of θ,

(ii) $\log(2\cos\theta)=\sum_{n=1}^{\infty}(-1)^{n-1}\dfrac{\cos 2n\theta}{n}$, $-\dfrac{\pi}{2}<\theta<\dfrac{\pi}{2}$;

(iii) $\frac{1}{2}\theta=\sum_{n=1}(-1)^{n-1}\dfrac{\sin n\theta}{}$ $\quad -\pi<\theta<\pi$;

or, $\pi-\theta$ being put in place of θ,

(iv) $\frac{1}{2}(\pi-\theta)=\sum_{n=1}^{\infty}\dfrac{\sin n\theta}{n}$, $0<\theta<2\pi$.

In (iii) the *value* of the series for $\theta=\pi$ or $-\pi$ is zero but the *limit* of the series when $\theta\to\pi$ is $\frac{1}{2}\pi$, and when $\theta\to-\pi$ is $-\frac{1}{2}\pi$.

In (iv) the *value* of the series for $\theta=0$ is 0 but the *limit* of the series when $\theta\to 0$ is $\frac{1}{2}\pi$.

8. Show that in the notation of § 74, II, $x=\varrho(\cos\theta+i\sin\theta)$, $\varrho<1$ and $-\pi<\theta\leqq\pi$, if $k=m$ a *real* number, the principal value of $(1+x)^m$

is $\quad (1+2\varrho\cos\theta+\varrho^2)^{\frac{m}{2}}(\cos m\varphi+i\sin m\varphi)$ where

$$\varphi=\tan^{-1}\{\varrho\sin\theta/(1+\varrho\cos\theta)\},\quad -\frac{\pi}{2}<\varphi<\frac{\pi}{2}.$$

9. Convergence of the binomial series for $(1+x)^m$ when m is real and $|x|=1$, $x=\cos\theta+i\sin\theta$ where $-\pi<\theta\leqq\pi$.

(i) Convergence absolute if $m>0$; (ii) convergence conditional if $0>m>-1$, and also θ not equal to π; (iii) divergence if $m\leqq-1$.

[Let a_n be the coefficient of x^n or $(\cos n\theta+i\sin n\theta)$; then

$$\frac{n+1}{m-n}\quad 1+\frac{m+1}{n}+\frac{A_n}{n^2};\ A_n \text{ bounded.}$$

If $m+1>1$ or $m>0$, the convergence is absolute.

If $m<0$, let $m=-\mu$, $\mu>0$; the term in x^n may now be written

$$b_n[\cos(n\theta+n\pi)+i\sin(n\theta+n\pi)],\ b_n=\frac{\mu(\mu+1)\dots(\mu+n-1)}{1.2\dots n};$$

$b_n\to 0$ when $n\to\infty$ if, and only if, $\mu<1$. Now apply Dirichlet's Test to the two real series.]

10. Deduce from Examples 8 and 9 that, when m is real and $|x|=1$, $x=\cos\theta+i\sin\theta$, $-\pi<\theta\leqq\pi$, the series

$$+\binom{m}{1}x+\binom{m}{2}x^2+\binom{m}{3}x^3+\dots$$

is equal to $(2\cos\frac{1}{2}\theta)^m(\cos\frac{1}{2}m\theta+i\sin\frac{1}{2}m\theta)$ for all values of m for which the series converges.

11. Show that the value of $P_n(x)$, defined in § 66, Ex. 5, holds when x is complex provided $|x|\leqq 1$.

12. If $x=\cos\theta$ (θ real), express $(1-2y\cos\theta+y^2)^{-\frac{1}{2}}$ in the form

$$(1-ye^{i\theta})^{-\frac{1}{2}}\times(1-ye^{-i\theta})^{-\frac{1}{2}},$$

expand each binomial in powers of y, find the product of the two series and show that

$$1+\sum_{n=1}^{\infty}P_n(\cos\theta)y^n=1+\sum_{n=1}^{\infty}u_n(\theta)y^n$$

where $u_n(\theta)$—which is equal to $P_n(\cos\theta)$—is equal to

$$\frac{(2n)!}{2^{2n}(n!)^2}\left\{2\cos n\theta+\frac{1.n}{1.(2n-1)}2\cos(n-2)\theta\right.$$
$$\left.+\frac{1.3}{1.2}\frac{n(n-1)}{(2n-1)(2n-3)}2\cos(n-4)\theta+\dots\right\}$$

where the series ends, if n is odd, with the term which contains $2\cos\theta$ as a factor, but, if n is even, with the term which contains $\cos(0.\theta)$, that is, unity as factor.

Since the coefficients are all positive $P_n(\cos\theta)$ has its greatest value when $\theta=0$ and then $P_n(\cos\theta)=1$ so that, when θ is real,

$$-1\leqq P_n(\cos\theta)\leqq 1.$$

CHAPTER VII

SUBSTITUTION OF A SERIES IN A SERIES. REVERSION OF SERIES. LAGRANGE'S EXPANSION. MAXIMA AND MINIMA OF FUNCTIONS OF SEVERAL VARIABLES

76. Power Series. The theorem of § 66 on the derangement of a series will now be applied to the expansion of functions in a power series.

Substitution of a Power Series in a Power Series.

Suppose the function $f(y)$ to be given as a power series, convergent for $|y| < s$,

$$f(y) = a_0 + a_1 y + a_2 y^2 + \ldots + a_m y^m + \ldots, \quad |y| < s \quad \ldots\ldots\ldots(1)$$

where y is a power series in x, convergent for $|x| < r$,

$$y = b_0 + b_1 x + b_2 x^2 + \ldots + b_n x^n + \ldots, \quad |x| < r. \quad \ldots\ldots\ldots(2)$$

If $f(y)$ were a series in powers of $(y - y_0)$ and $(y - y_0)$ a series in powers of $(x - x_0)$, these could be reduced to the forms in (1) and (2) by substituting y for $(y - y_0)$ and x for $(x - x_0)$ so that there is no loss of generality in using the given forms; this simplification of notation is frequently used.

The values of $y^2, y^3, \ldots y^m \ldots$ may be found as series in powers of x by the rule for multiplying series, applied to the series (2) and all these series converge for $|x| < r$. Now substitute for $y, y^2, \ldots$ in (1) and rearrange in powers of x; the rearrangement can be effected when the conditions of § 66 are satisfied. The series A_m of § 66 will take the form

$$A_m = a_m y^m = a_m(a_{m,0} + a_{m,1}x + a_{m,2}x^2 + \ldots + a_{m,n}x^n + \ldots) \ldots(3)$$

where the series in brackets is the mth power of the series (2). (For $m = 1$, $a_{m,0} = a_{1,0} = b_0, \ldots a_{m,n} = a_{1,n} = b_n$; for $m = 2$; $a_{2,0} = b_0^2$, $a_{2,1} = 2b_0 b_1$, $a_{2,2} = 2b_0 b_2 + b_1^2 \ldots$ and so on.)

Now the series given by A_m is absolutely convergent for $|x|<r$, but the conditions of § 66 demand that the series

$$\alpha_m = |a_m|(|a_{m,0}| + |a_{m,1}|\,|x| + |a_{m,2}|\,|x|^2 + \ldots + |a_{m,n}|\,|x|^n + \ldots)$$

should converge and also that $\Sigma\alpha_m$, not merely $\Sigma|A_m|$, should converge and a smaller value of $|x|$ may be needed to secure this, because α_m and $|A_m|$ are, as a rule, very different numbers.

We now use the symbol α_m in a different sense from that given to it in this reference to § 66.

Let $|a_n|=\alpha_n$, $|b_n|=\beta_n$ and $|x|=\xi$. We try to satisfy the conditions of § 66.

The first condition is plainly that $\beta_0<s$ because $y=b_0$ when $x=0$.

Again, the series (2), and therefore also the series (3), converges absolutely if $|x|\leqq\varrho<r$ so that $\beta_n\varrho^n$, being a term of a convergent series, is finite for every value of n, say $\beta_n\varrho^n<M$. From (2) we find

$$|y|\leqq\beta_0+\beta_1\xi+\beta_2\xi^2+\ldots+\beta_n\xi^n+\ldots$$

$$<\beta_0+M\left(\frac{\xi}{\varrho}+\frac{\xi^2}{\varrho^2}+\ldots+\frac{\xi^n}{\varrho^n}+\ldots\right)$$

and therefore $\quad |y|<\beta_0+M\xi/(\varrho-\xi)$.

If then ξ is chosen so that $\beta_0+M\xi/(\varrho-\xi)$ is less than s, that is, β_0 being less than s by the first condition, if ξ is such that

$$\xi<\frac{(s-\beta_0)\varrho}{s-\beta_0+M},\quad \beta_0<s \tag{4}$$

the series (1) will converge as required when the series (2) has been substituted in it for y and rearrangement is allowable. The series B_n of § 66 will be

$$B_n=\left(\sum_{m=0}^{\infty}a_m a_{m,n}\right)x^n=c_n x^n,\ \text{say},$$

and then

$$f(y)=\sum_{n=0}^{\infty}c_n x^n. \qquad .(5)$$

The substitution and rearrangement are therefore valid if

$$\text{(i)}\ \beta_0<s,\quad \text{(ii)}\ |x|<\frac{(s-\beta_0)\varrho}{s-\beta_0+M},\ \varrho<r \qquad \ldots\ldots(6)$$

where M is an upper limit to the values of $\beta_n\varrho^n$ and $\beta_n=|b_n|$.

Cor. 1. If $b_0=0$ only one condition is necessary, namely

$$|x|<s\rho/(s+M);$$

this case is of special importance because the coefficient of x^n in (5) contains only a finite number of terms and is not an infinite series.

Cor. 2. Again if $s=\infty$ there is only one condition,

$$|x|<\rho<r$$

and, since ρ may differ as little as we please from r, the transformation is valid simply if $|x|<r$.

In proofs of Existence Theorems it is usually the possibility and not the full range of a transformation that is in question; in the above case it is frequently possible to verify that the range of the variable x may be greater than the inequalities (6) would allow. It may be noted further that conditions (6) are merely sufficient, not necessary.

Ex. 1. The expansion of log $(1+\sin x)$ in powers of x. (*E.T.* p. 398.)

Here $$f(y)=\log(1+y)=y-\tfrac{1}{2}y^2+\tfrac{1}{3}y^3-\tfrac{1}{4}y^4+\dots\ |y|<1 \quad\dots\dots\dots\dots(1)$$

where $$y=\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\dots\quad\dots\dots\dots\dots(2)$$

If $|x|=\rho$ and all the terms in (2) are made positive the series becomes $\sinh\rho$. Now $\sinh 0{\cdot}88=0{\cdot}998<1$ and the transformation is valid if $|x|<0{\cdot}88$.

Ex. 2. $f(y)=e^y$, $y=k\log(1+x)$.

In this case the series for $f(y)$ converges for every value of y while the series for y is convergent if $|x|<1$. Therefore the transformation holds if $|x|<1$. Further $b_0=0$ so that the coefficient of x^n is a polynomial. Show that

$$c_n=\frac{1}{n!}k(k-1)(k-2)\dots(k-n+1),$$

and, since $f(y)=(1+x)^k$, deduce the binomial theorem.

77. Division by a Series. If the quotient u/v is required where u is a polynomial or an infinite series in powers of x, and v is an infinite series in powers of x we may first express $1/v$ as an infinite series and then find the product of u and $1/v$ by multiplication of series.

If $v=b_0+b_1x+b_2x^2+\dots$ we may suppose $b_0=1$ when b_0 is not zero; for b_0 may be taken out as a factor of the series and then b_n written in place of b_n/b_0. There are two cases.

(i) $b_0=1$. Let $$y=b_1x+b_2x^2+\ldots \mid x \mid < r \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$
then $$1/v=1/(1+y)=1-y+y^2-y^3+\ldots \mid y \mid < 1. \quad \ldots\ldots\ldots(2)$$

By § 76, we may when $|x|$ is less than some number, ρ' say, substitute the series (1) in the series (2) and rearrange in powers of x; the product of u and $1/v$ will then be of the form $\Sigma c_n x^n$.

(ii) Suppose $b_0=0$ and let the lowest power of x that occurs in v be x^p, its coefficient being taken to be unity, say

$$v=x^p+b_{p+1}x^{p+1}+b_{p+2}x^{p+2}+\ldots \mid x \mid < r \quad \ldots\ldots\ldots\ldots(3)$$
$$=x^p(1+b_{p+1}x+b_{p+2}x^2+\ldots).$$

Express $(1+b_{p+1}x+b_{p+2}x^2+\ldots)^{-1}$ as a power series $\Sigma c_n x^n$; then

$$\frac{u}{v}=\frac{1}{x^p}\times u\times(\Sigma c_n x^n)=\frac{1}{x^p}\sum_{n=0} d_n x^n,$$

so that $$\frac{u}{v}=\frac{d_0}{x^p}+\frac{d_1}{x^{p-1}}+\ldots+\frac{d_{p-1}}{x}+d_p+d_{p+1}x+d_{p+2}x^2+\ldots.$$

In practice it is usually simpler, now that the validity of the transformation is established, to apply the method of undetermined multipliers (*E.T.* p. 388, Ex. 10.)

Ex. Expansion of $x/(e^x-1)$.

$$e^x-1=\sum_{n=1}^{\infty}\frac{x^n}{n!} \quad \text{and therefore} \quad \frac{x}{e^x-1}=\frac{1}{v} \quad \text{where}$$

$$v=1+\frac{x}{2}+\frac{x^2}{3!}+\frac{x^3}{4!}+\ldots+\frac{x^n}{(n+1)!}+\ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

and the value of $x/(e^x-1)$ for $x=0$ is taken to be unity.

Now express $1/v$ in the form

$$1/v=c_0+c_1x+c_2x^2+\ldots+c_nx^n+\ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

where c_0, c_1, ... have to be determined. Multiply the series in (2) by the series for v in (1); then the following equation

$$1=\left(1+\frac{x}{2}+\frac{x^2}{3!}+\ldots\right)(c_0+c_1x+c_2x^2+\ldots) \quad \ldots\ldots\ldots\ldots\ldots(3)$$

must be identically true for $|x|<\rho$, where ρ is not definitely known except that it must be positive, not zero. Hence the absolute term c_0 on the right of (3) must be equal to 1, the only term on the left; while the coefficient of each power of x on the right of (3), when the multiplication has been effected, must be zero. Thus we find

$$c_0=1;\ \frac{c_0}{2}+c_1=0;\ \frac{c_0}{3!}+\frac{c_1}{2}+c_2=0;\ \frac{c_0}{4!}+\frac{c_1}{3!}+\frac{c_2}{2}+c_3=0$$

and, in general,

$$\frac{c_0}{(n+1)!}+\frac{c_1}{n!}+\frac{c_2}{(n-1)!}+\ldots+\frac{c_{n-1}}{2}+c_n=0.$$

These equations, when solved, give:

$$c_0=1,\ c_1=-\tfrac{1}{2},\ c_2=\tfrac{1}{12},\ c_3=0,\ c_4=-\tfrac{1}{720},\ c_5=0,\ \ldots.$$

It is hopeless to seek by this method the value of c_n except for the smaller values of n; it may be proved, however, that no odd power of x above the first occurs in the expansion. For, if we write

$$f(x)=\frac{x}{e^x-1}+\tfrac{1}{2}x,$$

we find that $f(-x)=f(x)$ so that $f(x)$ is an even function and contains only even powers of x. The fact, however, that there is a power series has been established.

See further § 94.

78. Reversion of Series. If y is defined as a function of x by the convergent series

$$y=a_1x+a_2x^2+\ldots+a_nx^n+\ldots,\ |x|<r,\ \ldots\ldots\ldots\ldots\text{(A)}$$

the problem of reversing the series in (A) is, in general terms, that of expressing x as a convergent series in powers of y. It has been pointed out in § 76 that there is no loss of generality in taking x, y instead of $x-x_0$, $y-y_0$ as the variables, and further, that no importance attaches to the particular value of r (provided r is not zero). In the following discussion, therefore, the essential point is that the *series converge*; the determination of the *maximum* range of convergence of the various series is a separate problem.

It is, however, desirable to reduce the equation (A) to a standard form before defining more precisely the problem of reversion.

Suppose in the first place that the coefficient a_1 is not zero and make the substitutions:

$$y/a_1=y',\quad a_n/a_1=a'_n,\quad n>1.$$

Equation (A) becomes

$$y'=x+a'_2x^2+a'_3x^3+\ldots+a'_nx^n+\ldots.\ \ldots\ldots\ldots\ldots(\text{A}_1)$$

Suppose next that the coefficient a_m is not zero but that $a_1, a_2, \ldots, a_{m-1}$ are all zero and make the substitutions:

$$y/a_m=y',\quad a_{m+p}/a_m=a'_{m+p},\quad p=1, 2, \ldots.$$

In this case equation (A) takes the form

$$y'=x^m+a'_{m+1}x^{m+1}+a'_{m+2}x^{m+2}+\ldots.\ \ldots\ldots\ldots\ldots(\text{A}_2)$$

The series in (A_1) and (A_2) will still be convergent; the fact that the coefficient of x in (A_1) and of x^m in (A_2) is unity is an

important simplification. The accents may now be dropped and the discussion will proceed on the basis of equations (A_1) and (A_2), as thus changed. The two equations require separate consideration.

First Case. Coefficient a_1 in (A) not zero. The equation to be considered is

$$y = x + a_2x^2 + a_3x^3 + \dots + a_nx^n + \dots . \quad \text{..............(1)}$$

The problem of reversing the series in (1) may now be stated as follows: to show that there is one, and only one, convergent series for x in powers of y, say

$$x = y + b_2y^2 + b_3y^3 + \dots + b_ny^n + \dots \quad \text{...............(2)}$$

which satisfies the two conditions that $x=0$ when $y=0$ and makes the equation (1) an identity when the series (2) is substituted for x in the equation (1).

We must first show that when $|x|$ is sufficiently small, y, as determined by equation (1), cannot be zero unless $x=0$. For,

$$y = x(1 + a_2x + a_3x^2 + \dots) = x(1+v) \text{ say.}$$

The series v, that is, $a_2x + a_3x^2 + \dots$, is convergent and therefore defines a continuous function of x (*E.T.* p. 386); further $v=0$ when $x=0$, and therefore, by the continuity of v, it is possible to choose $|x|$ so small, say $|x| < r_1$, as to make $|v| < 1$. Hence $1+v$ is positive if $|x| < r_1$ and therefore $x(1+v)$ is, if $|x| < r_1$, zero if, and only if, $x=0$.

Let it be assumed for the moment that there is at least one convergent series which when substituted for x in (1) makes that equation an identity. Since the coefficient of y in equation (1) is unity its coefficient in the series for x must also be unity; the equation (2) may therefore be taken as defining the assumed series for x. The solution of the problem of reversion consists in showing that this assumption is justifiable.

When the series (2) is substituted for x in (1) and the coefficients of $y^2, y^3, \dots$ calculated, each coefficient must be zero because the equation is then an identity; thus, the following equations connecting the a's and the b's are obtained.

$$\begin{aligned} &b_2 = -a_2, \quad b_3 = -a_2(2b_2) - a_3, \\ &b_4 = -a_2(b_2^2 + 2b_3) - a_3(3b_2) - a_4, \quad \text{................(3)} \\ &b_5 = \dots, \quad b_6 = \dots . \end{aligned}$$

These equations determine, in succession, $b_2, b_3, b_4, \ldots$ *uniquely* as polynomials in $a_2, a_3, a_4, \ldots$, the coefficients of the polynomials being positive or negative integers; that is, b_n is of the form,

$$b_n = P_n(a_2, a_3, \ldots, a_n).$$

(If the coefficient of x in (1) were not unity P_n would be a polynomial, not in $a_2, a_3, \ldots, a_n$ simply, but in $a_1, a_2, \ldots, a_n$, divided by a power of a_1.) This determination of the coefficients b_n proves the important result that if there is a series of the kind assumed there is *only one* such series.

The next step is, by a method due, like so much in the theory, to Cauchy, to solve a particular case of the problem and then by means of this solution to pass to that of the given problem.

Suppose that the series (1) converges for $|x| < r$; then it converges absolutely for $|x| \leqq \varrho < r$ and therefore there is a positive number M such that $|a_n|\varrho^n < M$ for every value of the integer n. Let $\alpha_n = M/\varrho^n$ and consider the problem for the particular case, where for distinction ξ, η are used in place of x, y respectively:

$$\eta = \xi - \alpha_2\xi^2 - \alpha_3\xi^3 - \ldots \quad |\xi| < \varrho < r \quad \text{.........(1a)}$$

$$\xi = \eta + \beta_2\eta^2 + \beta_3\eta^3 + \ldots . \quad \text{.........(2a)}$$

The equations corresponding to (3) are

$$\beta_2 = \alpha_2, \quad \beta_3 = \alpha_2(2\beta_2) + \alpha_3,$$
$$\beta_4 = \alpha_2(\beta_2^2 + 2\beta_3) + \alpha_3(3\beta_2) + \alpha_4, \quad \text{.........(3a)}$$
$$\beta_5 = \ldots, \quad \beta_6 = \ldots,$$

and therefore, since each α is positive, so is each β.

Now $\quad |a_n| < \alpha_n$ and therefore

$$|b_2| = |a_2| < \alpha_2; \quad |b_2| < \beta_2.$$
$$|b_3| \leqq |a_2|(2|b_2|) + |a_3| < \alpha_2(2\beta_2) + \alpha_3; \quad |b_3| < \beta_3,$$

and so on, $|b_n| < \beta_n$, $n = 2, 3, 4, \ldots$.

If then the series (2a) converges, say for $|\eta| < s$, the series (2) will converge for $|y| < s$ and then by § 76, Cor. 1, the substitution in (1) of the series (2) for x will be justified; the equation (1) will be identically satisfied and the problem of the reversion of the series (1) will be solved.

We now determine the series ($2a$). In ($1a$) put M/ϱ^n for α_n then if $|\xi| < \varrho$ we find, as another form of ($1a$),

$$\eta = \xi - M\left(\frac{\xi^2}{\varrho^2} + \frac{\xi^3}{\varrho^3} + \dots\right) = \xi - \frac{M\xi^2}{\varrho(\varrho - \xi)} \quad \dots(1a)$$

so that $$(M+\varrho)\xi^2 - (\varrho+\eta)\varrho\xi + \varrho^2\eta = 0.$$

Solving this quadratic for ξ and taking the negative sign of the square root since $\xi = 0$ when $\eta = 0$ we obtain the equation:

$$2(M+\varrho)\xi = \varrho(\varrho+\eta) - \varrho\{\varrho^2 - (4M+2\varrho)\eta + \eta^2\}^{\frac{1}{2}}.$$

Let $\varrho^2 = s_1 s_2$ and $4M + 2\varrho = s_1 + s_2$; s_1 and s_2 are positive and we take $s_1 < s_2$. Thus

$$2(M+\varrho)\xi = \varrho(\varrho+\eta) - \varrho^2(1-\eta/s_1)^{\frac{1}{2}}(1-\eta/s_2)^{\frac{1}{2}}.$$

If $|\eta| < s_1 < s_2$ the binomials $(1-\eta/s_1)^{\frac{1}{2}}$ and $(1-\eta/s_2)^{\frac{1}{2}}$ may be expanded in convergent series of powers of η and when these series are multiplied we obtain the equation:

$$2(M+\varrho)\xi = \varrho(\varrho+\eta) - \varrho^2\left\{1 - \tfrac{1}{2}\frac{s_1+s_2}{s_1 s_2}\eta - c_2\eta^2 - c_3 n^3 - \dots\right\}$$
$$= \varrho(\varrho+\eta) - \varrho^2\left\{1 - \frac{2M+\varrho}{\varrho^2}\eta - c_2\eta^2 - c_3\eta^3 - \dots\right\}$$

so that $$\xi = \eta + \frac{c_2\varrho^2}{2(M+\varrho)}\eta^2 + \frac{c_3\varrho^2}{2(M+\varrho)}\eta^3 + \dots \quad \dots\dots\dots\dots(4)$$

(There is no purpose to be served by evaluating $c_2, c_3, \dots$ in terms of s_1 and s_2 since the series is known to be convergent.)

It has been already pointed out that if there is one convergent series of the type ($2a$) there is only one; the series given by (4) is convergent and therefore the series given by (4) and by ($2a$) must be identical. Hence the series (2) is convergent and therefore solves the problem of the reversion of the series (1).

Ex. If in the series of Equation (1) the signs are alternately + and − so that

$$y = x - a_2x^2 + a_3x^3 - a_4x^4 + a_5x^5 - a_6x^6 + \dots,$$

show that x is given by

$$x = y - b_2y^2 + b_3y^3 - b_4y^4 + b_5y^5 - b_6y^6 + \dots$$

where $b_2, b_3, b_4, \dots$ are the same as in Equation (2).

Second Case. The coefficients $a_1, a_2, \dots, a_{m-1}$ in (A) zero, a_m not zero.

In this case the form of equation (A) becomes (A_2) or

$$y = x^m + a_{m+1}x^{m+1} + a_{m+2}x^{m+2} + \quad \text{...............(5)}$$
$$= x^m(1+w), \quad w = a_{m+1}x + a_{m+2}x^2 + \dots .$$

When $|x|$ is sufficiently small it may be proved as before that $1+w$ is positive so that $y=0$ only if $x=0$.

Let $y=\eta^m$; then η has m different values, given by

$$\eta = y^{\frac{1}{m}}\left(\cos\frac{2k\pi}{m} + i\sin\frac{2k\pi}{m}\right) = y^{\frac{1}{m}}\,\theta_k, \quad k=0,\ 1, \dots (m-1),$$

where $y^{1/m}$ is the principal value of the mth root of y. Hence, by taking the mth root and expanding $(1+w)^{1/m}$ in powers of x by the binomial theorem, we find that equation (5) may be represented by m equations of the type

$$y^{\frac{1}{m}}\,\theta_k = \eta = x(1 + c_1x + c_2x^2 + \dots) = x + c_1x^2 + c_2x^3 + \dots \quad \text{...(6)}$$

The work in the First Case is not essentially altered if y is complex, and therefore to each of the equations of the type (6) there corresponds an equation of the type

$$x = \eta + d_1\eta^2 + d_2\eta^3 + \dots , \quad \text{....................(7)}$$

and the different equations of the type (7) are obtained by putting η equal to $y^{1/m}\theta_k$.

Thus in this case there are m different series each of which is zero when $y=0$ and when substituted in (5) reduces it to an identity.

79. Lagrange's Expansion. In the equation

$$z = x + yf(z), \quad \text{..........................(1)}$$

let x be considered as a constant (or a parameter) and y as a function of z. If $f(z)$ can be expressed as a convergent series in powers of $(z-x)$, say

$$f(z) = a_0 + a_1(z-x) + a_2(z-x)^2 + \dots \quad \text{...............(2)}$$

where $a_0 = f(x)$ *and is not zero*, y may, by expressing $1/f(z)$ as a series in powers of $(z-x)$, be represented by a convergent series of the form

$$y = b_1(z-x) + b_2(z-x)^2 + \dots \quad b_1 = 1/a_0 \quad \text{............(3)}$$

and then, by reversion of this series,

$$z - x = c_1y + c_2y^2 + c_3y^3 + \dots \quad c_1 = 1/b_1 = a_0. \quad \text{..........(4)}$$

Again, if $\varphi(z)$ may be represented by a convergent power series in $(z-x)$ the substitution of the series (4) for $(z-x)$ will give for $\varphi(z)$ a convergent series of the form

$$\varphi(z)=d_0+d_1y+d_2y^2+\dots \qquad \text{.....................(5)}$$

When $f(z)$ and $\varphi(z)$ satisfy the conditions stated both $f(z)$ and $\varphi(z)$ possess nth derivatives with respect to y for all values of n and therefore the series in (5), which, by the conditions satisfied by $f(z)$ and $\varphi(z)$, is unique must be the same as Taylor's expansion of $\varphi(z)$ in powers of y.

The range of y for which the series (5) converges cannot, as a rule, be found by the theorems at our disposal but it is certain from the theory of reversion of series that the series does converge for $|y|<s$, where s is positive. A rule for determining s in certain cases will be stated at the end of the article.

The calculation of the derivatives of $\varphi(z)$ can be effected by a method due to Lagrange by which the derivatives with respect to y are expressed in terms of derivatives with respect to the parameter x and when the coefficients d_n in (5) are expressed as derivatives with respect to x the expansion (5) is generally called **Lagrange's Expansion** of $\varphi(z)$.

From the equation (1) we find

$$\frac{\partial z}{\partial x}=1+yf'(z)\frac{\partial z}{\partial x},\quad \frac{\partial z}{\partial y}=f(z)+yf'(z)\frac{\partial z}{\partial y},$$

and therefore, by eliminating $f'(z)$,

$$\frac{\partial z}{\partial y}=f(z)\frac{\partial z}{\partial x}. \qquad \text{..........................}(a)$$

Again, $$\frac{\partial\varphi(z)}{\partial y}=\varphi'(z)\frac{\partial z}{\partial y}=\varphi'(z)f(z)\frac{\partial z}{\partial x}. \qquad \text{.....................}(b)$$

Next let $\psi(z)$ be any differentiable function of z; then

$$\frac{\partial}{\partial y}\left\{\psi(z)\frac{\partial z}{\partial x}\right\}=\psi'(z)\frac{\partial z}{\partial y}\frac{\partial z}{\partial x}+\psi(z)\frac{\partial^2 z}{\partial y\partial x}=\frac{\partial}{\partial x}\left\{\psi(z)\frac{\partial z}{\partial y}\right\},$$

and therefore by (a)

$$\frac{\partial}{\partial y}\left\{\psi(z)\frac{\partial z}{\partial x}\right\}=\frac{\partial}{\partial x}\left\{\psi(z)f(z)\frac{\partial z}{\partial x}\right\}. \qquad \text{............................}(c)$$

Now let $\psi(z)=\varphi'(z)f(z)$; then by (b) and (c)

$$\frac{\partial^2\varphi(z)}{\partial y^2}=\frac{\partial}{\partial y}\left\{\varphi'(z)f(z)\frac{\partial z}{\partial x}\right\}=\frac{\partial}{\partial x}\left\{\varphi'(z)[f(z)]^2\frac{\partial z}{\partial x}\right\}. \quad \dots(c')$$

In this case the form of equation (A) becomes (A_2) or

$$y = x^m + a_{m+1}x^{m+1} + a_{m+2}x^{m+2} + \dots \quad \dots\dots\dots(5)$$
$$= x^m(1+w),\ w = a_{m+1}x + a_{m+2}x^2 + \dots .$$

When $|x|$ is sufficiently small it may be proved as before that $1+w$ is positive so that $y=0$ only if $x=0$.

Let $y=\eta^m$; then η has m different values, given by

$$\eta = y^{\frac{1}{m}}\left(\cos\frac{2k\pi}{m} + i\sin\frac{2k\pi}{m}\right) = y^{\frac{1}{m}}\,\theta_k,\ k=0,\ 1, \dots (m-1),$$

where $y^{1/m}$ is the principal value of the mth root of y. Hence, by taking the mth root and expanding $(1+w)^{1/m}$ in powers of x by the binomial theorem, we find that equation (5) may be represented by m equations of the type

$$y^{\frac{1}{m}}\,\theta_k = \eta = x(1 + c_1x + c_2x^2 + \dots) = x + c_1x^2 + c_2x^3 + \dots . \quad \dots(6)$$

The work in the First Case is not essentially altered if y is complex, and therefore to each of the equations of the type (6) there corresponds an equation of the type

$$x = \eta + d_1\eta^2 + d_2\eta^3 + \dots , \quad \dots\dots\dots(7)$$

and the different equations of the type (7) are obtained by putting η equal to $y^{1/m}\theta_k$.

Thus in this case there are m different series each of which is zero when $y=0$ and when substituted in (5) reduces it to an identity.

79. Lagrange's Expansion. In the equation

$$z = x + yf(z), \quad \dots\dots\dots(1)$$

let x be considered as a constant (or a parameter) and y as a function of z. If $f(z)$ can be expressed as a convergent series in powers of $(z-x)$, say

$$f(z) = a_0 + a_1(z-x) + a_2(z-x)^2 + \dots \quad \dots\dots\dots(2)$$

where $a_0 = f(x)$ *and is not zero*, y may, by expressing $1/f(z)$ as a series in powers of $(z-x)$, be represented by a convergent series of the form

$$y = b_1(z-x) + b_2(z-x)^2 + \dots \quad b_1 = 1/a_0 \quad \dots\dots\dots(3)$$

and then, by reversion of this series,

$$z - x = c_1y + c_2y^2 + c_3y^3 + \dots \quad c_1 = 1/b_1 = a_0. \quad \dots\dots\dots(4)$$

Again, if $\varphi(z)$ may be represented by a convergent power series in $(z-x)$ the substitution of the series (4) for $(z-x)$ will give for $\varphi(z)$ a convergent series of the form

$$\varphi(z)=d_0+d_1y+d_2y^2+\dots \quad\dots\dots\dots\dots\dots(5)$$

When $f(z)$ and $\varphi(z)$ satisfy the conditions stated both $f(z)$ and $\varphi(z)$ possess nth derivatives with respect to y for all values of n and therefore the series in (5), which, by the conditions satisfied by $f(z)$ and $\varphi(z)$, is unique must be the same as Taylor's expansion of $\varphi(z)$ in powers of y.

The range of y for which the series (5) converges cannot, as a rule, be found by the theorems at our disposal but it is certain from the theory of reversion of series that the series does converge for $|y|<s$, where s is positive. A rule for determining s in certain cases will be stated at the end of the article.

The calculation of the derivatives of $\varphi(z)$ can be effected by a method due to Lagrange by which the derivatives with respect to y are expressed in terms of derivatives with respect to the parameter x and when the coefficients d_n in (5) are expressed as derivatives with respect to x the expansion (5) is generally called **Lagrange's Expansion** of $\varphi(z)$.

From the equation (1) we find

$$\frac{\partial z}{\partial x}=1+yf'(z)\frac{\partial z}{\partial x},\quad \frac{\partial z}{\partial y}=f(z)+yf'(z)\frac{\partial z}{\partial y},$$

and therefore, by eliminating $f'(z)$,

$$\frac{\partial z}{\partial y}=f(z)\frac{\partial z}{\partial x}. \qquad .(a)$$

Again, $$\frac{\partial\varphi(z)}{\partial y}=\varphi'(z)\frac{\partial z}{\partial y}=\varphi'(z)f(z)\frac{\partial z}{\partial x}. \quad\dots\dots\dots\dots\dots .(b)$$

Next let $\psi(z)$ be any differentiable function of z; then

$$\frac{\partial}{\partial y}\left\{\psi(z)\frac{\partial z}{\partial x}\right\}=\psi'(z)\frac{\partial z}{\partial y}\frac{\partial z}{\partial x}+\psi(z)\frac{\partial^2 z}{\partial y\partial x}=\frac{\partial}{\partial x}\left\{\psi(z)\frac{\partial z}{\partial y}\right\},$$

and therefore by (a)

$$\frac{\partial}{\partial y}\left\{\psi(z)\frac{\partial z}{\partial x}\right\}=\frac{\partial}{\partial x}\left\{\psi(z)f(z)\frac{\partial z}{\partial x}\right\}. \quad\dots\dots\dots\dots\dots (c)$$

Now let $\psi(z)=\varphi'(z)f(z)$; then by (b) and (c)

$$\frac{\partial^2\varphi(z)}{\partial y^2}=\frac{\partial}{\partial y}\left\{\varphi'(z)f(z)\frac{\partial z}{\partial x}\right\}=\frac{\partial}{\partial x}\left\{\varphi'(z)[f(z)]^2\frac{\partial z}{\partial x}\right\}. \quad\dots(c')$$

Similarly by putting $\varphi'(z)[f(z)]^2$ for $\psi(z)$ in (c) we find

$$\frac{\partial^3\varphi(z)}{\partial y^3}=\frac{\partial}{\partial x}\cdot\frac{\partial}{\partial y}\left\{\varphi'(z)[f(z)]^2\frac{\partial z}{\partial x}\right\}=\frac{\partial^2}{\partial x^2}\left\{\varphi'(z)[f(z)]^3\frac{\partial z}{\partial x}\right\}. \quad \ldots(c'')$$

Let it be now assumed that the law suggested by (c') and (c'') is general, that is, that

$$\frac{\partial^n\varphi(z)}{\partial y^n}=\frac{\partial^{n-1}}{\partial x^{n-1}}\left\{\varphi'(z)[f(z)]^n\frac{\partial z}{\partial x}\right\}. \quad \ldots\ldots\ldots\ldots\ldots(d)$$

After putting $\varphi'(z)[f(z)]^n$ for $\psi(z)$ in (c), differentiate with respect to y and we obtain the equation

$$\frac{\partial^{n+1}\varphi(z)}{\partial y^{n+1}}=\frac{\partial^{n-1}}{\partial x^{n-1}}\cdot\frac{\partial}{\partial y}\left\{\varphi'(z)[f(z)]^n\frac{\partial z}{\partial x}\right\}=\frac{\partial^n}{\partial x^n}\left\{\varphi'(z)[f(z)]^{n+1}\frac{\partial z}{\partial x}\right\},$$

so that the $(n+1)$th derivative of $\varphi(z)$ with respect to y is of the same form as the nth derivative. Hence the form given by (d) holds for all values of n greater than unity; the form (b) may be treated as the *zero*th derivative of $\varphi'(z)f(z)\partial z/\partial x$ so that the law holds if $n>0$.

The variables x and y in these differentiations are independent. The value of the nth derivative of $\varphi(z)$ with respect to y for any given value of y may therefore be obtained by substituting that value in the right-hand member of equation (d) *either before or after* the differentiations with respect to x have been made. If the given value of y is zero then $z=x$ and $\partial z/\partial x=1$ so that

$$\left[\frac{\partial^n\varphi(z)}{\partial y^n}\right]_{y=0}=\frac{d^{n-1}}{dx^{n-1}}\{\varphi'(x)[f(x)]^n\}.$$

Now expand $\varphi(z)$ by Maclaurin's Theorem:

$$\begin{aligned}\varphi(z)&=[\varphi(z)]_{y=0}+y\left[\frac{\partial\varphi(z)}{\partial y}\right]_{y=0}+\ldots+\frac{y^n}{n!}\left[\frac{\partial^n\varphi(z)}{\partial y^n}\right]_{y=0}+\ldots\\&=\varphi(x)+y\varphi'(x)f(x)+\frac{y^2}{2!}\frac{d}{dx}\{\varphi'(x)[f(x)]^2\}+\\&\qquad+\frac{y^n}{n!}\frac{d^{n-1}}{dx^{n-1}}\{\varphi'(x)[f(x)]^n\}+ \qquad .(\mathrm{I})\end{aligned}$$

where the form d/dx may be used since $\varphi'(x)[f(x)]^n$ is a function of x alone.

The special case in which $\varphi(z)=z$ gives the expansion

$$z=x+yf(x)+\frac{y^2}{2!}\frac{d}{dx}\cdot[f(x)]^2+\ldots+\frac{y^n}{n!}\frac{d^{n-1}}{dx^{n-1}}[f(x)]^n+\ldots \qquad (\mathrm{II})$$

If $x=0$ equation (I) takes the form

$$\varphi(z)=\varphi(0)+y[\varphi'(x)f(x)] \quad +\frac{y^2}{2!}\left[\frac{d}{dx}\{\varphi'(x)[f(x)]^2\}\right]_{x=0}+\ldots$$
$$+\frac{y^n}{n!}\left[\frac{d^{n-1}}{dx^{n-1}}\{\varphi'(x)[f(x)]^n\}\right]_{x=0}+\ldots . \quad \ldots\ldots\ldots\ldots\text{(III)}$$

In all cases z is the value which satisfies the condition that $z=0$ when $y=0$.

For another form of the Expansion, see § 80, Ex. 13.

Note. Lagrange's Expansion requires for a complete discussion the theory of functions of a complex variable. Reference may be made to MacRobert, *Functions of a Complex Variable*, § 54, or Whittaker and Watson, *Modern Analysis*, § 7·32.

Hermite's *Cours, rédigé en* 1882 *par M. Andoyer*, contains on pages 182-197 a valuable discussion of Kepler's Equation (see Example 5, § 80). Serret's *Algèbre Supérieure*, 6th Ed. Vol. I. pp. 466-484, gives an exposition based on memoirs of Cauchy and Rouché.

An excellent presentation of Lagrange's method is given by Bromwich, *Infinite Series*, 2nd Ed., pp. 158-160, and pp. 265-266.

A rule for determining the range of y for which Lagrange's Expansion converges when $z=yf(z)$ may be stated as follows:

Let $|z|=r$ and let $\psi(r)$ be the least value of $|y|$, that is, of $|z\div f(z)|$ when $|z|=r$. If s is the maximum value of $\psi(r)$ Lagrange's Expansion converges when $|y|<s$.

If the equation is $z=x+yf(z)$ let $z=\zeta+x$ so that

$$\zeta=yf(\zeta+x)=yF(\zeta)$$

and proceed as before.

The rule is by no means evident but the application of it in the examples of § 80 makes its meaning clear; there are many cases, however, such as Example 5, in which the determination of s is laborious.

The proof of the rule depends on the theory of functions of the complex variable; see Bromwich, *l.c.* p. 265, or Goursat, *Cours*, Vol. II, pp. 123-4.

Serret (*l.c.* p. 480) states the rule in the form: If s is the least value of $|y|$ for which the equations

$$z = x + yf(z) \text{ and } 1 = yf'(z)$$

have a common root—that is, for which the equation $z = x + yf(z)$ has two equal roots in z—Lagrange's Expansion converges when $|y| < s$.

In whatever way the range of convergence of the series in (I), (II) and (III) may be determined, the corresponding expansions are valid for that range.

80. Examples. The following examples furnish illustrations of Lagrange's Expansion; some of them are worked out in full to indicate the general method of solution. The numbers (I), (II) and (III) refer to the equations of § 79.

Ex. 1. If $y = z(1+z)^m$ where m is a positive integer expand in powers of y that value of z which is zero when $y = 0$.

Write the equation in the form $z = y(1+z)^{-m}$; then $x = 0$, $f(z) = (1+z)^{-m}$, $\varphi(z) = z$. Hence $[f(x)]^n = (1+x)^{-mn}$; when the derivatives of $[f(x)]^n$ have been calculated the value 0 is to be put for x, and equation (II) gives

$$z = y - \frac{2m}{2!}y^2 + \frac{3m(3m+1)}{3!}y^3 - \frac{4m(4m+1)(4m+2)}{4!}y^4 + .$$

$$+ (-1)^{n-1}\frac{nm(nm+1)\dots(nm+n-2)}{n!}y^n + \dots .$$

The least value of $|y|$ when $|z| = r > 0$ is $r(1-r)^m$, and the maximum value of $r(1-r)^m$ is easily found to be $m^m/(m+1)^{m+1}$; the series just found converges if $|y|$ is less than this number.

Ex. 2. $y = z - az^{m+1}$, m a positive integer; find the series for that value of z which is zero when y is zero.

Proceed as in Ex. 1. $f(z) = (1-az^m)^{-1}$ so that $[f(x)]^n$ is $(1-ax^m)^{-n}$. The only derivatives of the powers of $f(x)$ that are not zero when $x = 0$ are the mth, $(2m)$th, $(3m)$th, ..., (nm)th ... and therefore the only powers of y that occur in the series after y itself are the $(m+1)$th, $(2m+1)$th, ... The values of the derivatives are easily found by expanding $(1-ax^m)^{-n}$ by the binomial theorem. Hence

$$z = y + ay^{m+1} + \frac{2m+2}{2}a^2y^{2m+1} + \dots$$

$$+ \frac{(nm+2)(nm+3)\dots(nm+n)}{n!}a^ny^{nm+1} + \dots .$$

Again, the least value of $|y|$ when $|z| = r$ is $r - Ar^{m+1}$ $(A = |a|)$, and the series converges if

$$|y| < \frac{m}{m+1}\{(m+1)A\}^{-\frac{1}{m}}.$$

Ex. 3. $z = x + yz^{m+1}$; $z = x$ for $y = 0$, m a positive integer and x positive.

(i) $z = x + yx^{m+1} + \dfrac{2m+2}{2!}y^2x^{2m+1} + \dfrac{(3m+3)(3m+2)}{3!}y^3x^{3m+1} + \dots$

(ii) $\log z = \log x + yx^m + \dfrac{2m+1}{2!}y^2x^{2m} + \dfrac{(3m+2)(3m+1)}{3!}y^3x^{3m} + \dots$

Both series converge if $|y| < m^m/(m+1)^{m+1}x^m$.

Ex. 4. Expand e^{az} in powers of y when $y = ze^{bz}$.

Here $\qquad z = ye^{-bz}$ so that $f(x) = e^{-bx}$, $\varphi(x) = e^{ax}$

and $\qquad \varphi'(x)[f(x)]^n = ae^{(a-nb)x}$. Equation (III) gives

$$e^{az} = 1 + ay + \frac{a(a-2b)}{2!}y^2 + \dots + \frac{a(a-nb)^{n-1}}{n!}y^n + \dots$$

and the series converges if $|by| < 1/e$.

Cor. If $a = 1$, $b = -1$, $x = e^z$ then $\log x = xy$, and we find

$$x = 1 + y + \frac{3}{2}y^2 + \frac{4^2}{3!}y^3 + \frac{5^3}{4!}y^4 + \dots + \frac{(n+1)^{n-1}}{n!}y^n + \dots .$$

Ex. 5. $z = x + y \sin z$. (Kepler's Equation.)

The determination of the general term in the expansion is a matter of difficulty and the student should consult the section in Hermite's *Cours* (see § 79, *Note*). It is easy enough to calculate a few of the earlier terms.

(i) $z = x + y\sin x + \dfrac{y^2}{2}\sin 2x + \dfrac{y^3}{8}(3\sin 3x - \sin x) + \dots .$

ii) $\sin z = \sin x + \dfrac{y}{2}\sin 2x + \dfrac{y^2}{8}(3\sin 3x - \sin x)$

$$+ \frac{y^3}{6}(2\sin 4x - \sin 2x) + \dots .$$

(iii) $1 - y\cos z = 1 - y\cos x + \dfrac{y^2}{2}(1 - \cos 2x) + \dfrac{3y^3}{8}(\cos x - \cos 3x)$

$$+ \frac{y^4}{3}(\cos 2x - \cos 4x) + \dots .$$

Ex. 6. If $z = 2 - y/z$ expand z^{-p} in powers of y, that value of z being taken which is equal to 2 when $y = 0$.

Here $f(x) = -x^{-1}$ and $\varphi(x) = x^{-p}$. Apply equation (I), putting 2 for x in the evaluated derivatives; the result is

$$z^{-p} = \frac{1}{2^p} + \frac{p}{2^p}\cdot\frac{y}{4} + \frac{p(p+3)}{2^p . 2!}\left(\frac{y}{4}\right)^2 + \frac{p(p+4)(p+5)}{2^p . 3!}\left(\frac{y}{4}\right)^3 + \dots ,$$

the general term being $\dfrac{p(p+n+1)(p+n+2)\dots(p+2n-1)}{2^p\, n!}\left(\dfrac{y}{4}\right)^n$,

and the series converges if $|y| < 1$.

Since $z = 1 + (1-y)^{\frac{1}{2}}$ the expansion may be stated in the form

$$\left\{\frac{1+(1-y)^{\frac{1}{2}}}{2}\right\}^{-p} = 1 + p\left(\frac{y}{4}\right) + \frac{p(p+3)}{2!}\left(\frac{y}{4}\right)^2 + \dots .$$

Cor. Let $(1-y)^{\frac{1}{2}}=1-2u$, $p=-n$; then

$$(1-u)^n=1-\frac{n}{1}u(1-u)+\frac{n(n-3)}{2!}u^2(1-u)^2-\frac{n(n-4)(n-5)}{3!}u^3(1-u)^3+\dots$$

where $$-\tfrac{1}{2}(\sqrt{2}-1)<u<\tfrac{1}{2}.$$

Ex. 7. In Ex. 6 let $y=4t$, where $|4t|<1$, and let $p=-k$ where k is a positive integer; show that if $P(t)$ is given by the equation

$$P(t)=\left\{\frac{1+(1-4t)^{\frac{1}{2}}}{2}\right\}^k+\left\{\frac{1-(1-4t)^{\frac{1}{2}}}{2}\right\}^k,$$

$P(t)$ is a polynomial of degree $\frac{1}{2}k$ or $\frac{1}{2}(k-1)$ according as k is even or odd. Then prove that

$$P(t)=1-kt+\frac{k(k-3)}{2!}t^2-\frac{k(k-4)(k-5)}{3!}t^3+\dots.$$

Note that $$\left\{\frac{1-(1-4t)^{\frac{1}{2}}}{2}\right\}^k=t^k\left\{\frac{1+(1-4t)^{\frac{1}{2}}}{2}\right\}^{-k},$$

so that the expansion of this part of $P(t)$ will cancel all except the terms of the polynomial that arise from the first part of $P(t)$.

Ex. 8. Prove that if $|4t|<1$

$$\log\left\{\frac{1-(1-4t)^{\frac{1}{2}}}{2t}\right\}=t+\tfrac{3}{2}t^2+\frac{4\,.\,5}{3!}t^3+\dots,$$

the general term being $\dfrac{(n+1)(n+2)\dots(2n-1)}{n!}t^n$.

Ex. 9. **Theorem.** *The derivatives of $\varphi(z)$ with respect to x are obtained by differentiating Lagrange's Expansion term by term with respect to x.*

$\varphi(z)$, by hypothesis, can be expressed as a convergent series $\Sigma c_n(z-x)^n$ and the derivatives $\partial^m\varphi(z)/\partial z^m$ are obtained by differentiating the series term by term (*E.T.* p. 400). But the form of the series shows that $\partial^m\varphi(z)/\partial x^m$ is simply

$$(-1)^m\partial^m\varphi(z)/\partial z^m$$

so that $\partial^m\varphi(z)/\partial x^m$ can be expressed as a convergent series in powers of $z-x$.

Now, by equation (d) of § 79,

$$\frac{\partial^n}{\partial y^n}\cdot\frac{\partial^m\varphi(z)}{\partial x^m}=\frac{\partial^m}{\partial x^m}\frac{\partial^n\varphi(z)}{\partial y^n}=\frac{\partial^m}{\partial x^m}\cdot\frac{\partial^{n-1}}{\partial x^{n-1}}\left\{\varphi'(z)[f(z)]^n\frac{\partial z}{\partial x}\right\},$$

and therefore

$$\left[\frac{\partial^n}{\partial y^n}\cdot\frac{\partial^m\varphi(z)}{\partial x^m}\right]_{y=0}=\frac{d^{m+n-1}}{dx^{m+n-1}}\left\{\varphi'(x)[f(x)]^n\right\},$$

so that the expansion of $\partial^m\varphi(z)/\partial x^m$ by Maclaurin's Theorem is

$$\frac{d^m\varphi(x)}{dx^m}+y\frac{d^m}{dx^m}\left\{\varphi'(x)f(x)\right\}+\ldots+\frac{y^n}{n!}\frac{d^{m+n-1}}{dx^{m+n-1}}\left\{\varphi'(x)[f(x)]^n\right\}+\ldots,$$

and this is the series obtained by differentiating the series for $\varphi(z)$ term by term.

Ex. 10. Prove that if m is a positive integer and $|y|<1$,

$$\frac{(x-y)^m}{(1-y)^{m+1}}=x^m+\sum_{n=1}^{\infty}\frac{y^n}{n!}\frac{d^n}{dx^n}\left\{x^m(x-1)^n\right\}. \qquad \text{(Serret.)}$$

Let $z=x+y(z-1)$ so that $f(z)=z-1$ and in the series for $\partial\varphi(z)/\partial x$ let $\varphi'(z)=z^m$.

Ex. 11. If $z=x+y\left(\frac{z^2-1}{2}\right)$, expand in powers of y that value of z which is equal to x when $y=0$.

Here $f(x)=\frac{1}{2}(x^2-1)$ and equation (II) gives at once

$$z=x+y\left(\frac{x^2-1}{2}\right)+\frac{y^2}{2!}\frac{d}{dx}\cdot\left(\frac{x^2-1}{2}\right)^2+\ldots+\frac{y^n}{n!}\frac{d^{n-1}}{dx^{n-1}}\cdot\left(\frac{x^2-1}{2}\right)^n+\ldots. \quad (1)$$

If we solve the quadratic equation for z, choosing the *negative* sign of the root, since $z=x$ when $y=0$, we find

$$z=\{1-(1-2xy+y^2)^{\frac{1}{2}}\}/y,$$

and therefore, differentiating z with respect to x,

$$\frac{\partial z}{\partial x}=(1-2xy+y^2)^{-\frac{1}{2}}=1+\sum_{n=1}^{\infty}P_n(x)y^n \quad\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

by § 66, Ex. 5, $P_n(x)$ being Legendre's Polynomial.

Now differentiate (1) with respect to x; therefore by Ex. 9,

$$\frac{\partial z}{\partial x}=1+\sum_{n=1}^{\infty}\frac{y^n}{n!}\frac{d^n}{dx^n}\cdot\left(\frac{x^2-1}{2}\right)^n, \quad\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

so that by equating coefficients of y^n in (2) and (3) we find

$$P_n(x)=\frac{1}{2^n\,.\,n!}\frac{d^n}{dx^n}\cdot(x^2-1)^n. \quad\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Equation (4) gives **Rodrigues' Formula** for $P_n(x)$.

Ex. 12. By differentiating the equation

$$v=(1-2xy+y^2)^{-\frac{1}{2}}=1+\sum_{n=1}^{\infty}P_n(x)y^n$$

with respect to y, show that

$$(x-y)(1-2xy+y^2)^{-\frac{1}{2}}=(1-2xy+y^2)\sum_{n=1}^{\infty}nP_n(x)y^{n-1},$$

and, by equating coefficients of y^n, prove that

$$(n+1)P_{n+1}(x)-(2n+1)xP_n(x)+nP_{n-1}(x)=0,\ n=0,\ 1,\ 2,\ \ldots\ \ldots\text{(i)}$$

where $P_{-1}(x)$ is taken to be unity.

Again by differentiating with respect to x, show that

$$(x-y)\frac{\partial v}{\partial x}=y\frac{\partial v}{\partial y},$$

that is, if $P'_n(x)$ denote $dP_n(x)/dx$,

$$(x-y)\sum_{n=1}^{\infty}P_n(x)y^n=y\sum_{n=1}^{\infty}nP_n(x)y^{n-1},$$

and deduce, by equating coefficients of y^n, that

$$xP'_n(x)-P'_{n-1}(x)=nP_n(x),\ n=1,\,2,\,\ldots\,. \qquad \text{(ii)}$$

Ex. 13. Show that

$$\text{(i)}\ \frac{\varphi'(z)f(z)}{1-yf'(z)}=\varphi'(x)f(x)+\sum_{n=1}^{\infty}\frac{y^n}{n!}\frac{d^n}{dx^n}\left\{\varphi'(x)[f(x)]^{n+1}\right\}$$

and, if $\psi(z)$ can be expressed as a convergent series $\Sigma c_n(z-x)^n$,

$$\text{(ii)}\ \frac{\psi(z)}{1-yf'(z)}=\psi(x)+\sum_{n=1}^{\infty}\frac{y^n}{n!}\frac{d^n}{dx^n}\left\{\psi(x)[f(x)]^n\right\}.$$

The expansion (i) is obtained by differentiating with respect to y the series (I) for $\varphi(z)$. The expansion (ii) is then found by putting $\psi(z)$ for $\varphi'(z)f(z)$; by the conditions for Lagrange's Expansion $\varphi'(z)f(z)$ can be expressed as a convergent series in powers of $(z-x)$.

Ex. 14. Prove that

$$\frac{1}{\sqrt{(1-2y-3y^2)}}=1+\sum_{n=1}^{\infty}\frac{y^n}{n!}\left[\frac{d^n\,.\,(1+x+x^2)^n}{dx^n}\right]_{x=0}$$

Apply Ex. 13 (ii) taking $f(z)=1+z+z^2$ and $\psi(z)=1$; the expansion begins with the terms

$$1+y+3y^2+7y^3+19y^4+\ldots\,.$$

81. Implicit Function of One Variable. It has been proved in Chapter V, § 52, that under certain conditions an equation $F(x, y)=0$ defines y as a function of x. We shall now consider a special case in which $F(x, y)$ is given by an infinite series and then sketch the proof of the corresponding general theorem.

Let M be a positive constant, u_n the homogeneous polynomial

$$u_n=x^n+x^{n-1}y+x^{n-2}y^2+\ldots+xy^{n-1}+y^n,$$

and

$$y=Mx+M\sum_{2}^{\infty}u_n \qquad (1)$$

the series in (1) being convergent, as will be proved, when $|x|<1$ and $|y|<1$. It is to be proved that if $y=0$ when $x=0$ the equation (1) defines y as a single-valued function of x.

To each side of equation (1) add $M+My$; the equation thus becomes, if $u_0=1$,

$$M+(M+1)y=M(u_0+u_1+u_2+\ldots+u_n+\ldots). \quad \ldots\ldots.(2)$$

The series in brackets contains every product of the type $x^m y^n$ where both m and n take every positive integral value, including zero. The terms in (2) that contain x^m form the series

$$Mx^m(1+y+y^2+\ldots+y^n+\ldots)$$

and the series formed by the moduli of the terms, namely,

$$M|x|^m(1+|y|+|y|^2+\ldots+|y|^n+\ldots)$$

converges if $|y|<1$, its sum being $M|x|^m(1-|y|)^{-1}$. Also

$$\sum_{m=0}^{\infty} M|x|^m(1-|y|)^{-1}=M(1-|x|)^{-1}(1-|y|)^{-1},$$

provided $|x|<1$. Hence the double series in (2) is convergent if $|x|<1$ and $|y|<1$ and its sum is therefore given by

$$M\sum_{m=0}^{\infty} x^m \sum_{n=0}^{\infty} y^n = M(1-x)^{-1}(1-y)^{-1}$$

so that $$M+(M+1)y=M(1-x)^{-1}(1-y)^{-1}.$$

This equation is a quadratic in y,

$$(M+1)y^2-y+Mx(1-x)^{-1}=0 \quad \ldots\ldots\ldots\ldots\ldots.(3)$$

and, when solved for y, gives

$$2(M+1)y=1-(1-x)^{-\frac{1}{2}}\{1-(2M+1)^2x\}^{\frac{1}{2}} \quad \ldots\ldots\ldots..(4)$$

where the negative value of the root has been chosen so that we may have $y=0$ when $x=0$.

If $|x|<1/(2M+1)^2$ (and therefore also $|x|<1$), each binomial can be expanded in a convergent series of powers of x and, when the series have been multiplied, y will be given by a convergent series

$$y=Mx+c_2x^2+c_3x^3+\ldots. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots.(5)$$

The theorem is therefore proved. The general theorem of which this is a special case may be stated as follows:

THEOREM. *Let u_n denote the polynomial*

$$u_n=a_{n,0}\,x^n+a_{n-1,1}\,x^{n-1}y+a_{n-2,2}\,x^{n-2}y^2+\ldots+a_{0,n}\,y^n,$$

and let $$F(x,y)=-y+a_{1,0}\,x+\sum_{n=2}^{\infty} u_n \quad \ldots\ldots\ldots\ldots\ldots..(\mathrm{I})$$

where the coefficient of y is -1 and the series converges for

$|x|<1$ *and* $|y|<1$. *The equation* $F(x, y)=0$ *defines* y *as a single-valued function of* x *by a convergent series, say*

$$y=b_1x+b_2x^2+b_3x^3+\dots+b_nx^n+\dots, \qquad \text{(II)}$$

which satisfies the condition that $y=0$ *when* $x=0$. *The series, when substituted for* y *in* $F(x, y)$, *makes* $F(x, y)$ *identically zero.*

The coefficient of the first power of y in $F(x, y)$ must not be zero, and there is no loss of generality in assigning the stated form to $F(x, y)$. If the series converges for $|x|<R_1$ and $|y|<R_2$ put $x=R_1x'$, $y=R_2y'$, and if the coefficient of y' is not -1 but k, say, divide the equation $F(R_1x', R_2y')=0$ by $-k$; the new form of the function is that assumed in the above statement. When the transformation has been completed the accents may be dropped from x', y' and the coefficients denoted by the symbols given. Again $F(x, y)$ *may be a polynomial*, that is, after a certain stage each coefficient $a_{m,n}$ may be zero.

The method of proof follows the lines of that used in the theorem of the Reversion of Series. Suppose, to begin with, that the coefficients b_n in (II) are undetermined; substitute the series for y in $F(x, y)$, and, if possible, choose $b_1, b_2, \dots$ so that when $F(x, y)$ is arranged in powers of x the coefficient of each power will be zero. If this choice can be made, and if the values of $b_1, b_2, \dots$ so found are unique, the condition that $F(x, y)$ vanishes identically will be *formally* satisfied, and the theorem will be *formally* proved since $y=0$ when $x=0$. To make the proof complete (that is, a *real* proof) it must be shown that the series (II), with the values of $b_1, b_2, \dots$ that have been found, is a *convergent* series; when it is convergent the various transformations are valid.

Now, the equating of coefficients determines $b_1, b_2, \dots$ in succession (compare § 78) and b_n is given by a polynomial

$$b_n=P_n(a_{1,0}, a_{2,0}, a_{1,1}, a_{0,2}, \dots a_{0,n}) \qquad \text{(III)}$$

in which the coefficients are *positive integers*. This determination is unique and therefore if there is one convergent series such as (II) there is only one.

Next, the general term in the series (I) is $a_{m,n}x^my^n$, and since the series converges for $|x|<1$ and $|y|<1$ there is a positive number, M say, such that $|a_{m,n}|<M$ for every value of m and n. Take now the equation $\varphi(\xi, \eta)=0$ where

$$\varphi(\xi, \eta)=-\eta+M\xi+M(\xi^2+\xi\eta+\eta^2+\xi^3+\xi^2\eta+\dots) \qquad \text{(Ia)}$$

and let
$$\eta=\beta_1\xi+\beta_2\xi^2+\beta_3\xi^3+\dots+\beta_n\xi^n+\dots. \qquad \text{(IIa)}$$

Substitute this value of η in $\varphi(\xi, \eta)$ and choose $\beta_1, \beta_2, \ldots$ so that $\varphi(\xi, \eta)$ may be identically zero. The value of β_n is given by the polynomial in (III) when M has been substituted for each of the numbers $a_{1,0}, a_{2,0}, \ldots$; since the coefficients in P_n are positive integers and $|a_{m,n}| < M$ the number β_n is *positive and greater than* $|b_n|$. It has, however, been proved that equation (5) with ξ, η in place of x, y gives a convergent series which makes $\varphi(\xi, \eta)$ identically zero and therefore the series in (5) and (IIa) must be the same. Hence the series (IIa) and therefore the series (II), since $|b_n| < \beta_n$, converge, so that the proof of the theorem is now complete.

Factorisation. In $F(x, y)$ substitute $P(x)+z$ for y, where $P(x)$ is the series in (II), and arrange as a series in powers of x and z; the function $F[x, P(x)+z]$ is identically zero when $z=0$ and therefore $F[x, P(x)+z]$ is of the form $zP_1(x, z)$ where $P_1(x, z)$ is a series in powers of x and z. Let z be now replaced by $y-P(x)$ and we find

$$F(x, y) = [y - P(x)]\, P_2(x, y)$$

where $P_2(x, y)$ is a series in powers of x and y, the absolute term being -1 because the coefficient of y in $F(x, y)$ is -1. The analogy with the usual expression $f(x) = (x-a) f_1(x)$ when $f(a)=0$ is obvious. This Factorisation Theorem is due to Weierstrass.

Cor. If $F(x_0, y_0)=0$, the substitution $x=x_0+x'$ and $y=y_0+y'$ reduces the problem of finding a series for y in powers of x which is such that $y=y_0$ when $x=x_0$ to the problem just discussed for the function $F(x_0+x', y_0+y')$ or $F_1(x', y')$.

For applications and extensions of the above theorem the student is referred to Chrystal's *Algebra*, Part II, Chap. XXX, pp. 373-397. Some illustrations are given in Chapter XII of the *Elementary Treatise* (§§ 106, 107, and Exercises XX).

82. Algebraic Forms. As a preliminary to the consideration of the Remainder in Taylor's Theorem for a function of several variables it is necessary to prove two theorems on the behaviour of the ratio of two algebraic forms. The number of variables that appear in the statements will usually be three, x, y, z, but the definitions and the theorems are quite general. It is to

be understood that the variables and constants are all *real* numbers.

A polynomial that is homogeneous and of the nth degree in two or more variables is called a **Form** or a **Quantic**. A form $f(x, y, z)$ is a continuous function of its variables and therefore (§ 43, Th. II) if it has opposite signs for the values a, b, c and a', b', c' respectively of x, y, z—or, as will be often said, at the points (a, b, c) and (a', b', c')—it will be zero for an unlimited number of values of x, y, z, or at an unlimited number of points (x, y, z).

A form is said to be *definite* if it is not zero unless its variables are all zero and to be *indefinite* if it is zero for values of its variables that are not all zero.

For example, $x^4 + y^4 + z^4$ is a definite form. Every form of odd degree in the variables is indefinite because in that case $f(-x, -y, -z)$ and $f(x, y, z)$ have opposite signs for all values of x, y, z (not all zero).

A *definite form has the same sign* for all values of its variables (unless these are all zero) because, as has just been seen, if it had opposite signs at two points it would be zero at an unlimited number of points. The form is called a positive definite form or a negative definite form according as the sign is positive or negative.

It is possible, however, for a form to be neither definite nor indefinite; it may, like a definite form, have the same sign when it is not zero and *yet be zero when its variables are not all zero*. In this case the form is said to be *semi-definite*. For example, the form $(x + 2y - z)^2$ is semi-definite; it is never negative but it is zero at all points in the plane $z = x + 2y$.

Let $Q = f(x, y, z)/g(x, y, z)$ where f and g are two forms of the same degree; though each form is defined and continuous for all values of the variables, Q is not defined for values of the variables that are all zero. The point $(0, 0, 0)$ is a limiting point of the region for which Q is defined but does not belong to it so that the region is not closed (§ 40). The proof that a continuous function attains its upper and lower bounds, however, requires that its region of definition should be closed and, as the proof is important for the applications to be made

of the properties of forms, it will be shown how a closed region may be obtained.

Let $\qquad x=r\xi,\ y=r\eta,\ z=r\zeta,\ r=|(x^2+y^2+z^2)^{\frac{1}{2}}|$;

then $\qquad Q=f(\xi,\eta,\zeta)/g(\xi,\eta,\zeta)$,

where $\qquad \xi^2+\eta^2+\zeta^2=1. \quad \text{..............................(S)}$

Q is not defined for $x=0$, $y=0$, $z=0$, and r is not zero unless x, y, z are all zero; all the values for which Q is defined may thus be obtained by assigning to ξ, η, ζ values which satisfy equation (S). Now the region defined by (S), which for three variables is the surface of a sphere, contains its limiting points; for, if $P(a, b, c)$ is a limiting point of a set $P'(a+h, b+k, c+l)$ which lies on (S) then

$$(a+h)^2+(b+k)^2+(c+l)^2=1,$$

and therefore when h, k, l all tend to zero the numbers a, b, c satisfy equation (S) so that P lies on (S). In other words the region defined by equation (S) contains all its limiting points and is therefore closed; the reasoning is clearly applicable to the case of n variables $x, y, z, w, \ldots$.

Now equation (S) is not satisfied when ξ, η, ζ are all zero, and therefore when $g(x, y, z)$ is a *definite* form $g(\xi, \eta, \zeta)$ cannot be zero for any admissible values of ξ, η, ζ.

Two theorems will now be proved which are essential for the discussion to be given of the Remainder in Taylor's Theorem and these, with the whole discussion of the Remainder, are based on the exposition in the *Calculus* of Genocchi-Peano (German Translation, pp. 170-189).

THEOREM I. *If $f(x, y, z)$ and $g(x, y, z)$ are forms of degree n the ratio $f(x, y, z)/g(x, y, z)$ has an upper bound M and a lower bound m which are attained and are therefore maximum and minimum values of the ratio, provided $g(x, y, z)$ is a definite form.*

Let the ratio be transformed in the manner just shown to $f(\xi, \eta, \zeta)/g(\xi, \eta, \zeta)$. The ratio is a continuous function of ξ, η, ζ, since $g(\xi, \eta, \zeta)$ is not zero at any point (ξ, η, ζ), and therefore has upper and lower bounds, M and m respectively, which are attained and are therefore maximum and minimum values of the ratio.

THEOREM II. *Let $g(x, y, z)$ be, as in Theorem I, a definite form of degree n and let $f(x, y, z)$ be expressible as the sum*

$$c_1(x, y, z) f_1(x, y, z) + c_2(x, y, z) f_2(x, y, z) + \ldots + c_m(x, y, z) f_m(x, y, z)$$

where $f_1, f_2, \ldots, f_m$ are forms of degree n while the coefficients $c_1, c_2, \ldots, c_m$ are functions of x, y, z, each of which tends to zero when all the variables tend to zero. The ratio $f(x, y, z)/g(x, y, z)$ tends to zero when all the variables x, y, z tend to zero.

For
$$\frac{f(x, y, z)}{g(x, y, z)} = \sum_{r=1}^{m} c_r(x, y, z) \frac{f_r(x, y, z)}{g(x, y, z)}$$

Each ratio f_r/g is bounded, by Theorem I, and each of the coefficients c_r tends to zero when x, y, z all tend to zero so that f/g also tends to zero.

Ex. 1. Find the condition that the form $ax^2 + 2bxy + cy^2$ should be definite.

If $y = 0$ the form becomes ax^2 so that a cannot be zero if the form is definite; similarly c cannot be zero and must have the same sign as a. Again
$$ax^2 + 2bxy + cy^2 = a(x + by/a)^2 + (ac - b^2)y^2/a$$
so that if $ac > b^2$ the form has the same sign as that of a (or c) and is therefore definite.

If $ac = b^2$ the form is semi-definite. If $ac < b^2$ the factors of the form are real and different and the form is then indefinite.

Ex. 2. If $\varphi(x, y, z) = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$, the form φ is definite if (i) a, b, c are all of the same sign, (ii) A, B, C are all positive, where A, B, C are the co-factors of a, b, c in the discriminant Δ of the form, and (iii) Δ has the same sign as a (or b or c).

$\varphi(1, 0, 0) = a$, $\varphi(0, 1, 0) = b$, $\varphi(0, 0, 1) = c$ and therefore a, b, c must be all different from zero and have the same sign when the form is definite. Again
$$\varphi(x, y, z) = a\left(x + \frac{hy + gz}{a}\right)^2 + \frac{C}{a}\left(y - \frac{F}{C}z\right)^2 + \frac{\Delta}{C}z^2$$
and therefore C must be positive and Δ must not be zero and must have the same sign as a.

Also $a\Delta = BC - F^2$, $b\Delta = CA - G^2$, $c\Delta = AB - H^2$ (where F, G, H are the co-factors of f, g, h in Δ). But $a\Delta$ is positive and therefore BC, and therefore B, is positive; similarly A is positive.

In general, when a form of the second degree in any number of variables is given the terms in one variable, say x, are brought together (as has been done above) into one term $a(x + b'y + c'z + d'w + \ldots)^2$; the terms left give a form in which the number of variables has been reduced by one and this form is treated in a similar way. Finally, a form in one variable is left.

The student may consult treatises on Higher Algebra, such as Bôcher's

Introduction to Higher Algebra. Bromwich's *Quadratic Forms* is useful; see also Hilton's *Linear Substitutions.*

Ex. 3. Show that $(x^2-y^2)/(x^2+y^2)$ and $(x^2+y^2-z^2)/(x^2+y^2+z^2)$ have $+1$ as the maximum value and -1 as the minimum value.

Ex. 4. The forms

$$\text{(i)}\ xy+y^2+2yz+3zx+wy+wz,$$

$$\text{(ii)}\ xy+2yz+3zx+wy+wz\ ;$$

can be expressed as

$$AX_1^2+BX_2^2-CX_3^2-DX_4^2$$

where A, B, C, D are positive numbers and X_1, X_2, X_3, X_4 are linear functions of x, y, z, w.

83. Remainder in Taylor's Theorem. Suppose that $f(x, y, z)$ is continuous at (a, b, c) and can be expanded near (a, b, c) by Taylor's Theorem; let $x=a+ht$, $y=b+kt$, $z=c+lt$, where $|h|$, $|k|$ and $|l|$ are small, and $f(x, y, z)=F(t)$, $f(a, b, c)=F(0)$. In the notation of § 157 of the *Elementary Treatise*, when $t=1$ and therefore x, y, z equal to $a+h$, $b+k$, $c+l$ respectively and $f(x, y, z)=F(1)$, the expansion is given by the equation

$$F(1)=F(0)+F'(0)+\dots+\frac{1}{r!}F^{(r)}(0)+\dots+\frac{1}{n!}F^{(n)}(\theta)\quad\dots(1)$$

where $0<\theta<1$ and $F^{(n)}(\theta)/n!=R_n$, the Remainder after n terms.

$F^{(r)}(0)$ is a form of degree r in the variables h, k, l, namely

$$F^{(r)}(0)=\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}+l\frac{\partial}{\partial c}\right)^r f(a, b, c)$$

$$=h^r\frac{\partial^r f}{\partial a^r}+rh^{r-1}k\frac{\partial^r f}{\partial a^{r-1}\partial b}+\dots+l^r\frac{\partial^r f}{\partial c^r}\qquad\text{(A)}$$

$$=\Sigma A_{\alpha,\beta,\gamma}h^\alpha k^\beta l^\gamma,\quad \alpha+\beta+\gamma=r$$

where $A_{\alpha,\beta,\gamma}$ is a function of a, b, c. $F^{(n)}(\theta)$ is of degree n in h, k, l and the coefficients, $A'_{\alpha,\beta,\gamma}$ say, in this case are functions of $a+\theta h$, $b+\theta k$, $c+\theta l$.

The equation (1) may be written so as to contain an additional term with the remainder R_{n+1}, namely,

$$F(1)=u_0+u_1+\dots+u_r+\dots+u_n+R_{n+1}\ \dots\dots\dots\dots(2)$$

where $u_r=F^{(r)}(0)/r!$, $r=1, 2, \dots, n$, and $u_n+R_{n+1}=R_n$, so that

$$R_{n+1}=R_n-u_n=\{F^{(n)}(\theta)-F^{(n)}(0)\}/n!\ \dots\dots\dots\dots(3)$$

THEOREM I. *If the term u_n is a definite form in the variables h, k, l the ratio of R_{n+1} to u_n tends to zero, and the ratio of R_n to u_n tends to unity, when all the variables h, k, l tend to zero.*

The expression for R_{n+1} given by (3) shows that R_{n+1} is a form of degree n in the variables h, k, l, namely

$$R_{n+1} = \frac{1}{n!}\Sigma(A_{\alpha,\beta,\gamma} - A_{\alpha,\beta,\gamma})h^{\alpha}k^{\beta}l^{\gamma},\ \alpha+\beta+\gamma=n\ ;$$

but $(A'_{\alpha,\beta,\gamma} - A_{\alpha,\beta,\gamma}) \to 0$ when all the variables h, k, l tend to zero and therefore by Theorem II of § 82 the ratio of R_{n+1} to u_n tends to zero. Hence when all the variables tend to zero

$$\frac{R_{n+1}}{u_n} \to 0,\ \frac{R_n - u_n}{u_n} \to 0,\ \text{that is, } \frac{R_n}{u_n} \to 1.$$

It should be specially noted that the proof is essentially conditioned by the assumption that u_n is a *definite* form.

THEOREM II. *Suppose that $f(a, b, c)$ or $F(0)$ is zero and that the first of the terms in* (1) *or* (2) *that does not vanish identically is u_r or $F^{(r)}(0)/r!$ There are two cases.* (i) *If $F^{(r)}(0)$ is a definite form $f(x, y, z)$ is not zero and is always of the same sign in the neighbourhood of (a, b, c) (that is, when h, k, l are not all zero).* (ii) *If $F^{(r)}(0)$ is an indefinite form $f(x, y, z)$ takes positive, negative and zero values in the neighbourhood of (a, b, c).*

Case (i). In the equation (1) let $n=r$; then $f(x, y, z)$ or $F(1)$ is $F^{(r)}(\theta)/r!$ The form $F^{(r)}(0)$ is definite and therefore by Theorem I $F^{(r)}(\theta)/F^{(r)}(0)$ tends to unity; since $F^{(r)}(\theta)$ is continuous in h, k, l and $F^{(r)}(0)$ is not zero and is always of the same sign so is $F^{(r)}(\theta)$ and therefore $f(x, y, z)$.

Case (ii). In this case the form $F^{(r)}(0)$ is indefinite and therefore there are values of h, k, l for which it is positive and also values for which it is negative. But $F^{(r)}(\theta) \to F^{(r)}(0)$ when h, k and l all tend to zero, and therefore, since it tends to values that are positive or negative according to the choice of h, k, l it must itself take both positive and negative values. Further $F^{(r)}(\theta)$ is a continuous function of h, k and l and therefore must also take zero values.

If $f(a, b, c)$ and $g(a, b, c)$ are each zero the fraction $f(x, y, z)/g(x, y, z)$ is not defined for the values a, b, c of x, y, z and (compare *E.T.* § 161) $f(a, b, c)/g(a, b, c)$ used to be called an "Indeterminate Form." The following theorem

throws some light on the possibilities of what may happen when x, y, z tend to a, b, c respectively.

It is understood that $f(x, y, z)$ and $g(x, y, z)$ can be expressed near (a, b, c) by Taylor's Theorem. Let $G(t)$ have the same meaning for $g(x, y, z)$ as $F(t)$ has in the preceding theorems for $f(x, y, z)$; then $F(0)=f(a, b, c)=0$, $G(0)=g(a, b, c)=0$ and

$$F(1)=u_1+u_2+\ldots+u_r+\ldots$$
$$G(1)=v_1+v_2+\ldots+v_r+\ldots$$

where $u_r=F^{(r)}(0)/r!$ and $v_r=G^{(r)}(0)/r!$

Suppose that $F^{(r)}(0)$ vanishes identically, that is, for all values of h, k, l, when r has the values $1, 2, \ldots (m-1)$ but not when $r=m$ and that $G^{(r)}(0)$ vanishes identically when $r=1, 2, \ldots (n-1)$ but not when $r=n$. Further, let w denote the quotient $f(x, y, z)/g(x, y, z)$.

THEOREM III. *If $G^{(n)}(0)$ is a definite form the quotient w or $f(x,y,z)/g(x,y,z)$ behaves, when all the variables h, k, l tend to zero, in the way specified in the following cases:*

Case (i), $m>n$: w tends to zero;

Case (ii), $m=n$: w oscillates finitely unless the ratio $F^{(n)}(0)/G^{(n)}(0)$ tends to a limit, K say, in which case w also tends to K;

Case (iii), $m<n$: w does not tend to any finite limit.

Case (i). Let $g(x, y, z)=G(1)=G^{(n)}(\theta')/n!$, $0<\theta'<1$; then $G^{(n)}(\theta')/G^{(n)}(0)\to 1$ because $G^{(n)}(0)$ is a definite form (Theorem I).

Next let $f(x, y, z)=F(1)=F^{(n)}(\theta)/n!$ as in equation (1); though $F^{(n)}(0)$ is identically zero since $m>n$ the form $F^{(n)}(\theta)$ is not zero because $F(1)$ is not zero. When all the variables h, k, l tend to zero the coefficients $A'_{\alpha,\beta,\gamma}$ in the form $F^{(n)}(\theta)$ tend to the coefficients $A_{\alpha,\beta,\gamma}$ in the form $F^{(n)}(0)$ and therefore to zero since $F^{(n)}(0)$ is identically zero. Hence by Theorem II of § 82, since $G^{(n)}(0)$ is a definite form, the ratio $F^{(n)}(\theta)/G^{(n)}(0)$ tends to zero and therefore w also tends to zero when h, k, l tend to zero because

$$w=\frac{F^{(n)}(\theta)}{G^{(n)}(\theta')}=\frac{F^{(n)}(\theta)}{G^{(n)}(0)} : \frac{G^{(n)}(\theta')}{G^{(n)}(0)}$$

and $G^{(n)}(\theta')/G^{(n)}(0)$ tends to unity.

Case (ii), $m=n$. Let $F(1)=u_n+R_{n+1}$ where R_{n+1} is given by equation (3). By the same reasoning as before the ratio

of $(n!\,R_{n+1})$, that is, of $\{F^{(n)}(\theta)-F^{(n)}(0)\}$ to the definite form $G^{(n)}(0)$ tends to zero when all the variables h, k, l tend to zero and therefore the behaviour of the ratio $F^{(n)}(\theta)/G^{(n)}(0)$ when $h, k. l$ tend to zero is the same as that of the ratio $F^{(n)}(0)/G^{(n)}(0)$.

Now
$$w=\frac{F^{(n)}(\theta)}{G^{(n)}(\theta')}=\frac{F^{(n)}(\theta)}{G^{(n)}(0)}:\frac{G^{(n)}(\theta')}{G^{(n)}(0)},$$

and $G^{(n)}(\theta')/G^{(n)}(0)$ tends to unity so that w behaves in the same way as $F^{(n)}(0)/G^{(n)}(0)$ when all the variables h, k, l tend to zero. But, by Theorem I of § 82, the ratio $F^{(n)}(0)/G^{(n)}(0)$ has a maximum value M and a minimum value m and therefore w oscillates between M and m unless the ratio $F^{(n)}(0)/G^{(n)}(0)$ tends to a limit K in which case $M=m=K$ and then w tends to K.

Cases (iii) $m<n$. In this case take the functions $F(t)$ and $G(t)$ instead of $F(1)$ and $G(1)$; we now have

$$w=\frac{F(t)}{G(t)}=\frac{1}{t^{n-m}}\cdot\frac{n!}{m!}\,\frac{F^{(m)}(\theta t)}{G^{(n)}(\theta' t)}.$$

If the form $F^{(m)}(0)$ is definite $F^{(m)}(\theta t)\rightarrow F^{(m)}(0)$ when $t\rightarrow 0$ and $F^{(m)}(0)$ is not zero unless h, k, l are all zero. If $F^{(m)}(0)$ is indefinite h, k, l can be chosen so that $F^{(m)}(0)$ is not zero since $F^{(m)}(0)$ is not identically zero. Therefore whether $F^{(m)}(0)$ is definite or indefinite, the ratio of $F^{(m)}(\theta t)$ to $G^{(n)}(\theta' t)$, when $t\rightarrow 0$, may be made to tend to $F^{(m)}(0)/G^{(n)}(0)$, or N say, where N is not zero. Hence w cannot tend to a *finite* limit when $t\rightarrow 0$, since $m<n$ and therefore $1/t^{n-m}$ tends to infinity while the factor N is not zero.

84. Maxima and Minima. The difficulty noticed in § 159 of the *Elementary Treatise* can now be to a certain extent cleared up. When $f(x, y)$ is continuous near (a, b) the derivatives f_a and f_b are both zero and, in the notation of the preceding article, we have for two variables h, k,

$$f(a+h,\, b+k)-f(a,\, b)=F''(\theta)/2.$$

The conclusions of Theorem II of § 83 are now applied. If $F''(0)$ is a definite form $F''(\theta)$ is not zero and is always of the same sign in the neighbourhood of (a, b), that is, when h and k are not both zero. On the other hand, if $F''(0)$ is an indefinite form $F''(\theta)$ takes both positive and negative values in the neighbourhood of (a, b). Hence, when $F''(0)$ is a

definite form $f(a, b)$ is a maximum value of $f(x, y)$ when the form is negative and a minimum when the form is positive but if $F''(0)$ is an indefinite form $f(x, y)$ is neither a maximum nor a minimum for the values a, b of x, y.

It is seen in the same way that if $F^{(r)}(0)/r!$ is the first term in the expansion of $f(a+h, b+k) - f(a, b)$ that is not identically zero $f(x, y)$ will be a maximum or a minimum for $x=a$, $y=b$ if $F^{(r)}(0)$ is a definite form but will be neither a maximum nor a minimum if $F^{(r)}(0)$ is an indefinite form.

These conclusions obviously hold for functions of any number of variables.

Nothing has been said of what conclusion may be drawn when $F''(0)$ is a *semi-definite form*; this is the case of Peano's example (*E.T.* p. 413). All that can be said in this case is that the above tests for a maximum or a minimum fail and further examination is necessary to decide the question of a maximum or a minimum. In many of the cases that occur in ordinary work it is often possible, as with Peano's example, to decide the question by use of purely algebraic methods but any general method usually involves complicated expansions, and even then may not lead to a definite conclusion. See, for example, Jordan's *Cours d'Analyse*, Vol. I, §§ 399, 400; Stolz's *Differential- und Integral-Rechnung*, Vol. I, *Abschnitt* V. (with the references); Hobson's *Functions of a Real Variable*, Chapter VI.

Ex. $f(x, y) = x^4 + y^4 - 2(x-y)^2$.

The equations $f_x = 0$, $f_y = 0$ give the points $(\sqrt{2}, -\sqrt{2})$, $(-\sqrt{2}, \sqrt{2})$ and $(0, 0)$ as points for which $f(x, y)$ may have a maximum or a minimum value. The test from the character of the form $F''(0)$ shows that $f(x, y)$ is a minimum at $(\sqrt{2}, -\sqrt{2})$ and at $(-\sqrt{2}, \sqrt{2})$; the test fails for the point $(0, 0)$.

But $f(\lambda, \lambda) = 2\lambda^4 > 0$, $f(\lambda, 0) = \lambda^2(\lambda^2 - 2) < 0$ if $\lambda^2 < 2$; since $f(0, 0) = 0$ the function $f(x, y)$ takes both positive and negative values in the neighbourhood of $(0, 0)$, because $|\lambda|$ may be arbitrarily small, and therefore $f(x, y)$ has neither a maximum nor a minimum value at $(0, 0)$.

In this and similar cases the consideration of the surface $z = f(x, y)$ is often useful.

It may, however, be remarked that the determination of maximum and minimum values can frequently be effected by the use of algebraic inequalities, as noted in the *Elementary Treatise* (§ 76), and the discussion in Chrystal's *Algebra*, Vol. II,

Chapter XXIV, will repay careful study. The methods of the Calculus are powerful, but it is a great mistake to neglect the resources of comparatively simple and straightforward algebra.

85. Absolute Maxima and Minima. When a function, $f(x)$ say, is defined for the range $a \leqq x \leqq b$ it does not follow, of course, that even when more than one maximum value has been found the greatest of these is also the greatest value of $f(x)$ in the range. It may quite well happen that $f(a)$ or $f(b)$ or both $f(a)$ and $f(b)$ may be greater than any value of $f(x)$ for the range $a<x<b$; the method of the calculus implies that the values of x for which the function is a maximum or a minimum lie *inside the range.* To find the *absolute* maximum or minimum it is therefore necessary to find the turning values, as determined by the Calculus, and then to compare them with each other and with $f(a)$ and $f(b)$.

It may be the case that $f(a)$ and $f(b)$ are themselves the maximum and minimum values and that $f(x)$ has no turning value between a and b.

For example, the perpendicular p from the focus $(ae, 0)$ of the ellipse $x^2/a^2+y^2/b^2=1$ on the tangent at the point (x, y) is given by the equation

$$p=b(a-ex)^{\frac{1}{2}}(a+ex)^{-\frac{1}{2}}$$

and

$$\frac{dp}{dx}=\frac{-abe}{(a+ex)(a^2-e^2x^2)^{\frac{1}{2}}}$$

so that dp/dx is neither zero nor infinite in the range $(-a, a)$. As x increases from $-a$ to a the perpendicular p steadily decreases from $a(1+e)$ to $a(1-e)$, so that $a(1+e)$ is the absolute maximum and $a(1-e)$ the absolute minumum value of p.

Though the maximum and minimum values of p cannot be found by the ordinary rule yet *the sign of the derivative* settles the matter; even when the derivative is discontinuous the sign will often indicate the possibility of a maximum or a minimum. Thus the function $f(x)$ where

$$f(x)=a+b^{\frac{1}{3}}(x-c)^{\frac{2}{3}}$$

is a minimum when $x=c$; $f'(x)$ is discontinuous for $x=c$ but $f'(x)$ is negative when $x<c$ and positive when $x>c$, so that $f(c)$ is the minimum value of the function.

In the case of functions of more than one variable corre-

sponding observations may be made. The type of region defined by the equation $\xi^2+\eta^2+\zeta^2=1$ in § 82 is of importance in some connections, for example, in the problem there treated and in the case of the determinant presently to be considered. The region has no boundaries in the usual meaning of the term, and every maximum or minimum that may occur is given by values that are *within the region.*

Implicit Functions of one Variable. If the equation $f(x, y)=0$ defines y as a function of x, the determination of the turning values of y is of course a frequent problem in the tracing of the curve represented by the equation and hardly demands any special treatment.

Since
$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dy}{dx}=0 \quad \text{.........................(1)}$$
the condition that $dy/dx=0$ gives the equation $f_x=0$ and the two equations $f=0$, $f_x=0$ determine the possible values of x and y. The sign of d^2y/dx^2 has next to be considered; differentiating equation (1) and noting that $dy/dx=0$ for a turning value we get
$$\frac{\partial^2 f}{dx^2}+\frac{\partial f}{\partial y}\frac{d^2y}{dx^2}=0. \quad \text{.........................(2)}$$

What is the significance of the coefficient of dy/dx in (1) and of d^2y/dx^2 in equation (2)?

86. Hadamard's Determinant. Let D be a determinant of the nth order, the element in the rth row and sth column being a_{rs}; if the numbers a_{rs} satisfy the n conditions
$$\varphi_r \equiv a_{r1}^2+a_{r2}^2+\ldots+a_{rs}^2+\ldots+a_{rn}^2-b_r=0 \quad \text{...........(1)}$$
where r has the values $1, 2, \ldots, n$ and b_r is constant, then
$$|D| \leqq \sqrt{(b_1 b_2 \ldots b_n)}$$
when the elements a_{rs} vary continuously.

The region defined by the equations (1) has no boundaries and therefore D, being a continuous function of its elements a_{rs}, has *both a maximum and a minimum value* which may be obtained by the usual method of undetermined multipliers (*E.T.* pp. 414, 415). Let A_{rs} be the co-factor of a_{rs} in D; then
$$D=a_{r1}A_{r1}+\ldots+a_{rs}A_{rs}+\ldots+a_{rn}A_{rn},\ r=1, 2, \ldots, n. \ \text{...(2)}$$
Now take the multipliers $\frac{1}{2}\lambda_1, \frac{1}{2}\lambda_2, \ldots, \frac{1}{2}\lambda_n$ and let
$$F\equiv D-\tfrac{1}{2}\lambda_1\varphi_1-\tfrac{1}{2}\lambda_2\varphi_2-\ldots-\tfrac{1}{2}\lambda_r\varphi_r-\ldots-\tfrac{1}{2}\lambda_n\varphi_n, \ \text{......(3)}$$

and the values of a_{rs} that determine the turning values of D are obtained by equating to zero the differentials of the elements in F. The coefficient of a_{rs} in D is A_{rs}, and the only one of the functions φ in which a_{rs} occurs is φ_r; hence we find

$$A_{rs} - \lambda_r a_{rs} = 0, \quad r = 1, 2, \dots, n, \; s = 1, 2, \dots, n. \quad \dots\dots\dots(4)$$

From (4), keeping r constant, assigning to s the values $1, 2, \dots, n$ and multiplying by a_{rs} we obtain the equations

$$a_{r1}A_{r1} + a_{r2}A_{r2} + \dots + a_{rn}A_{rn} = \lambda_r(a_{r1}^2 + a_{r2}^2 + \dots + a_{rn}^2),$$

that is,
$$D = \lambda_r b_r, \dots \quad r = 1, 2, \dots n. \quad \dots\dots\dots(5)$$

Again, instead of the multiplier a_{rs} take a_{ts} where t is any of the integers $1, 2, \dots, n$ except r; then keeping r and t fixed we get the equation

$$a_{t1}A_{r1} + a_{t2}A_{r2} + \dots + a_{tn}A_{rn} = \lambda_r(a_{t1}a_{r1} + a_{t2}a_{r2} + \dots + a_{tn}a_{rn}),$$

that is,
$$a_{r1}a_{t1} + a_{r2}a_{t2} + \dots + a_{rn}a_{tn} = 0, \quad \dots\dots\dots(6)$$

since $a_{t1}A_{r1} + \dots + a_{tn}A_{rn} = 0$.

From equations (4) and (5)

$$A_{rs} = a_{rs}D/b_r. \quad \dots\dots\dots(7)$$

In the equations (6), r and t are any two different integers from 1 to n so that when D has its maximum or minimum value the determinant is *orthogonal*, that is, the sum of the n products of corresponding elements in any two rows (say the rth and the tth) is zero. (Compare the equations (1) and (6) with the relations between the direction-cosines of three mutually perpendicular lines.)

The determinant which has A_{rs} as the element in the rth row and sth column is equal to D^{n-1}; but that determinant is by (7) equal to $D^{n+1}/b_1b_2 \dots b_n$. Hence the maximum and the minimum values of D are given by the equation

$$D^{n+1} = b_1b_2 \dots b_n D^{n-1}, \text{ that is, } D^2 = b_1b_2 \dots b_n,$$

so that
$$|D| \leqq \sqrt{(b_1b_2 \dots b_n)}$$

whatever values the elements a_{rs} may take so long as they satisfy the conditions (1).

Cor. If $|a_{rs}| \leqq M$, and therefore $b_r \leqq nM^2$, then

$$|D| \leqq (n^n)^{\frac{1}{2}} M^n.$$

The above theorem is due to Hadamard and is of great importance in the theory of Integral Equations.

EXERCISES IX.

(For answers to some of the Examples, see at end of the Set.)

Find the maximum and minimum values of the functions in Examples 1-12:

1. $xy^2(3x+6y-2)$.
2. $y^2+4xy+3x^2+x^3$.
3. $(x-1)(y-1)(x^2+y^2-4)$.
4. $x^4+2x^2y-x^2+3y^2$.
5. $y^2+x^2y+ax^4$.
6. $y^2+2z^2-5x^4+4x^5$.
7. $x^4+y^4+z^4-4xyz$.
8. $xy+x^{-1}+y^{-1}$.
9. $(x+y-1)/(x^2+y^2)$.
10. $(x+y)/(x^2+2y^2+6)$.
11. $(x-y)/(x^2+y^2+1)$.
12. $(y+z)^2+(z+x)^2+xyz$.

13. If all the letters denote positive numbers, show that the maximum value of

$$xy(z-h)\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}-\right.$$

is $(2h/5)^5 ab/c^4$.

14. If $7x^3+30x^2y+21y^3=21$, find the maximum and minimum values of x^2+y^2.

15. The maximum value of $xyz/(a+x)(x+y)(y+z)(z+b)$, where all the letters denote positive numbers, is given by

$$\frac{x}{a}=\frac{y}{x}=\frac{z}{y}=\frac{b}{\ }\quad\left(\frac{b}{\ }\right)^{\frac{1}{4}}$$

16. If $3a^2y^2+xy^3+4ax^3=0$, show that y has a maximum value, $-3a$, when $x=\frac{3a}{2}$, and that, if $2x^5+3ay^4-x^2y^3=0$, y is a minimum, $a\,.\,5^{5/3}$, when $x=a\,.\,5^{4/3}$. (a is positive.) (Todhunter.)

17. If $xyz=8$, the product $(x+1)(y+1)(z+1)$ is a minimum when $x=y=z=2$, and if $xyz=b^3$ the minimum value of the product

$$(x^2+a^2)(y^2+a^2)(z^2+a^2)$$

is $(a^2+b^2)^3$.

18. If $xyz=k^3$, the product $(x+a)(y+b)(z+c)$ is a minimum when

$$\frac{x}{a}=\frac{y}{b}=\frac{z}{c}=\frac{k}{(abc)^{1/3}}$$

the letters denoting positive numbers.

In Examples 19-24 all the letters denote positive numbers.

19. If $f(x, y, z)=x^l y^m z^n$ and $g(x, y, z)=x+y+z$, apply the tests for discriminating maxima and minima to prove that f possesses a maximum when g is constant and that g possesses a minimum when f is constant.

Show that the theorem holds for p variables $x_1, x_2, \dots x_p$, and extend to the more general form

$$g(x, y, z)=ax^\alpha+by^\beta+cz^\gamma.$$

20. (i) If $2x+3y+4z=a$, the maximum value of $x^2y^3z^4$ is $(a/9)^9$.

(ii) If $a^2x^2+2by^3+z^4=c^4$, the maximum value of x^4yz^2 is given by

$$17a^2x^2=12c^4,\ 17by^3=c^4,\ 17z^4=3c^4.$$

21. If $xyz=abc$, the minimum value of $bcx+cay+abz$ is $3abc$.

22. If $x^2/a^2+y^2/b^2+z^2/c^2=1$, the maximum value of xyz is $abc/3\sqrt{3}$. Interpret geometrically.

23. If $xyz=a^2(x+y+z)$, the minimum value of $yz+zx+xy$ is $9a^2$.

24. If $x^2+y^2=1$, the minimum value of $(ax^2+by^2)/(a^2x^2+b^2y^2)^{\frac{1}{2}}$ is $2(ab)^{\frac{1}{2}}/(a+b)$.

25. $\varphi(x,y,z)\equiv ax^2+by^2+cz^2+2fyz+2gzx+2hxy$, and

$\psi(x,y,z)\equiv lx+my+nz$;

find the maximum and minimum values of r^2 where r^2 is equal to $x^2+y^2+z^2$

(i) if $\varphi=k=$constant; (ii) if $\varphi=k$ and $\psi=0$.

If $u=-k/r^2$, the values of r^2 are the roots of the equations

$$\text{(i)}\ \begin{vmatrix} a+u, & h, & g \\ h, & b+u, & f \\ g, & f, & c+u \end{vmatrix}=0;\qquad \text{(ii)}\ \begin{vmatrix} a+u, & h, & g, & l \\ h, & b+u, & f, & m \\ g, & f, & c+u, & n \\ l, & m, & n, & 0 \end{vmatrix}=0.$$

When $\varphi=k$ represents an ellipsoid the volume of the ellipsoid and the area of the section by the plane $\psi=0$ are $4\pi r_1r_2r_3/3$ and $\pi\varrho_1\varrho_2$ where r_1, r_2, r_3 are the roots of Equation (1) and ϱ_1, ϱ_2 those of Equation (ii); these products are easily found.

26. If $(x^2+y^2+z^2)^2=a^2x^2+b^2y^2+c^2z^2$ and $lx+my+nz=0$, show that the maximum and minimum values of $r^2(=x^2+y^2+z^2)$ are given by the equation

$$l^2/(r^2-a^2)+m^2/(r^2-b^2)+n^2/(r^2-c^2)=0.$$

27. If $f(x,y,z)=(a^2x^2+b^2y^2+c^2z^2)/x^2y^2z^2$, where $ax^2+by^2+cz^2=1$ and a, b, c are positive, show that the minimum value of $f(x,y,z)$ is given by

$$\frac{u}{2a(u+a)},\quad y^2=\frac{\quad}{2b(u+b)},\quad \frac{\quad}{2c(u+c)}$$

where u is the positive root of the equation

$$u^3-(bc+ca+ab)u-2abc=0. \qquad \text{(Schlömilch.)}$$

28. $f(x,y,z)$ and $g(x,y,z)$ are two quadratic forms

$$f\equiv a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{23}yz+2a_{31}zx+2a_{12}xy,$$
$$g\equiv b_{11}x^2+b_{22}y^2+b_{33}z^2+2b_{23}yz+2b_{31}zx+2b_{12}xy;$$

if g is a positive definite form, prove that the maximum and minimum values of the ratio f/g are the values of u given by the equation

$$\begin{matrix} a_{11}-b_{11}u, & a_{12}-b_{12}u, & a_{13}-b_{13}u \\ a_{21}-b_{21}u, & a_{22}-b_{22}u, & a_{23}-b_{23}u \\ a_{31}-b_{31}u, & a_{32}-b_{32}u, & a_{33}-b_{33}u \end{matrix}=0.$$

where $a_{rs}=a_{sr}$, $b_{rs}=b_{sr}$. $(r=1,2,3,\ s=1,2,3)$.

If $f=x^2+7z^2-2yz-2xy$, $g=x^2+2y^2+5z^2+2yz-2xy$, show that u has the values $2, 1, -1$.

29. If $f(x, y) = 8x^2 - 6xy^2 + y^4$, and if $x = ht$, $y = kt$, $f(x, y) = F(t)$, show that $F(t)$ is a minimum when $t = 0$ $(h \neq 0,\ k \neq 0)$ although (*E.T.* p. 413) $f(x, y)$ is not a minimum for $x = 0$, $y = 0$.

Hence if $x = a + ht$, $y = b + kt$, $f(x, y) = F(t)$ it is possible that $F(0)$ may be a turning value of $F(t)$ and yet $f(a, b)$ not a turning value of $f(x, y)$.

30. From the point $B(0, -b)$ of the ellipse $x^2/a^2 + y^2/b^2 = 1$ a chord BP of the ellipse is drawn ; find the position of P when the length of BP is greatest. (Lampe.)

1. Min. at $(\frac{1}{6}, \frac{1}{6})$. 2. Min. at $(\frac{2}{3}, -\frac{4}{3})$.

3. Two minima given by $2x^2 + 2y^2 = x + y + 4$ and $x = y$ and two maxima given by $2x^2 + 2y^2 = x + y + 4$ and $x + y = 1$.

4. Min. at $\left(\frac{\sqrt{3}}{2}, -\frac{1}{4}\right)$ and at $\left(-\frac{\sqrt{3}}{2}\ \ -\frac{1}{4}\right)$.

5. Min. at $(0, 0)$ if $a > \frac{1}{4}$. 6. Min. at $(1, 0, 0)$.

7. Min. at $(1, 1, 1)$. 8. Min. at $(1, 1)$.

9. Max. at $(1, 1)$. 10. Max. at $(2, 1)$, Min. at $(-2, -1)$.

11. Max. at $\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$, Min. at $\left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

12. No Maximum or Minimum.

14. The values are obtained from the given equation and the three equations $x = 0$, $y = 2x$, $y = 5x$; the first two give minimum values and the third gives a maximum.

30. If $a^2 > 2b^2$, P is given by $x = \pm a^2(a^2 - 2b^2)^{\frac{1}{2}}/(a^2 - b^2)$, $y = b^3/(a^2 - b^2)$, but if $a^2 < 2b^2$, P is $(0, b)$.

CHAPTER VIII

INFINITE PRODUCTS. PRODUCTS AND SERIES OF PARTIAL FRACTIONS FOR TRIGONOMETRIC FUNCTIONS. GAMMA FUNCTIONS

87. Infinite Products. Let $f_1, f_2, f_3, \ldots$ be a sequence of real or complex numbers and let P_n be the product $f_1 f_2 \ldots f_n$, or, in the usual symbols

$$P_n = \prod_{r=1}^{n} f_r.$$

Definition. If when n tends to infinity P_n tends to a limit, P say, *which is not zero unless one of the factors f_r is zero*, the infinite product is said "to converge" or "to be convergent" and P is called the value of the product or simply "the product." If P_n tends to $+\infty$ or to $-\infty$ or (when no factor f_r is zero) to zero the infinite product is said "to diverge" or "to be divergent." If P_n tends to no definite limit (finite or infinite) the infinite product is said "to oscillate."

It is possible for P_n to tend to zero even though no factor f_r is zero; for example, $P_n = 1/(n+1)$ when $f_r = r/(r+1)$ and $P_n \to 0$ when $n \to \infty$. Of course, if any one factor f_r is zero so is P_n when $n \geqq r$ and therefore P_n tends to zero; but, by considering products that tend to zero when no factor f_r is zero as divergent, the property that a product does not vanish unless one of its factors vanishes, remains for *convergent* products.

Thus, if no factor f_r is zero and if P_n tends to a limit P that is not zero, $|P_n|$ must be greater than a positive constant, C say, *for every value of n*; because m may be chosen so that $|P_n| > C_1 > 0$ when $n > m$, while $|P_n|$ is not zero when n takes any one of the m values, $1, 2, \ldots, m$, since no factor f_r is zero. Hence $|P_n| > C$ for every value of n, where C is any positive constant less than C_1 or than any of the m numbers $|P_1|$, $|P_2|$,

In testing for convergence the condition that the limit P is not zero unless a factor f_r is zero must be specially noted.

The notation for an infinite product is, in analogy with that for an infinite series,

$$\prod_{n=1}^{\infty} f_n \quad \text{or} \quad \Pi f_n \quad \text{or} \quad \prod_{n=1}^{\infty} (1+u_n \quad \text{or} \quad \Pi(1+u_n)$$

the brackets being used when the factor contains two or more terms.

In the following work the symbol ε has the usual meaning.

THEOREM. *The product P_n, where $P_n = f_1 f_2 \dots f_n$, will, when n tends to infinity, tend to a limit P, that is not zero unless one of the factors f_r is zero, provided there is an integer m such that*

(i) $|f_{n+1} f_{n+2} \cdots f_{n+p} - 1| < \varepsilon$ if $n \geqq m$, $p = 1, 2, 3, \dots$.

This condition is equivalent to the following which is often more convenient in practice:

(ii) $\underset{n \to \infty}{\mathcal{L}} (f_{n+1} f_{n+2} \cdots f_{n+p} - 1) = 0$, $p = 1, 2, 3, \dots$.

The product $f_{n+1} f_{n+2} \cdots f_{n+p}$ is P_{n+p}/P_n.

(*a*) The condition is necessary. If no factor f_r is zero $|P_n|$ is greater than a positive constant C for every value of n while, if P_n tends to a limit, m may be chosen so that $|P_{n+p} - P_n| < \varepsilon C$ when $n \geqq m$ whatever integer p may be. Hence

$$|f_{n+1} f_{n+2} \cdots f_{n+p} - 1| = \frac{|P_{n+p} - P_n|}{|P_n|} < \frac{\varepsilon C}{C}, \; n \geqq m,$$

so that $|f_{n+1} f_{n+2} \cdots f_{n+p} - 1| < \varepsilon$ if $n \geqq m$, $p = 1, 2, 3, \dots$. The condition is therefore necessary.

(*b*) The condition is sufficient. Let $\varepsilon = \frac{1}{2}$; since condition (i) is satisfied m may be chosen so that, whatever integer n may be provided that $n > m$,

$$|f_{m+1} f_{m+2} \cdots f_n - 1| < \tfrac{1}{2}, \quad \text{or,} \quad \tfrac{1}{2} < |f_{m+1} f_{m+2} \cdots f_n| < \tfrac{3}{2},$$

and therefore, m being now fixed,

$$\tfrac{1}{2}|P_m| < |P_n| < \tfrac{3}{2}|P_m| \quad \text{if } n \geqq m, \quad \dots\dots\dots\dots\dots \quad (k)$$

If therefore P_n tends to a limit that limit cannot be zero unless a factor of P_m is zero. We have now

$$|P_{n+p} - P_n| = |P_n| \; |f_{n+1} f_{n+2} \cdots f_{n+p} - 1|.$$

But $|P_n| < \frac{3}{2}|P_m| < K$, a constant, by (k); and, since condition (i) is satisfied, there is an integer μ (which may be

taken greater than m) such that $|f_{n+1}\ldots f_{n+p}-1|<\varepsilon/K$ when $n\geqq\mu$, $p=1, 2, 3, \ldots$. Hence

$$|P_{n+p}-P_n|<\varepsilon \quad\text{if}\quad n\geqq\mu,\ p=1, 2, 3, \ldots,$$

so that P_n tends to a limit which, as has been shown, is not zero.

Cor. 1. Let $p=1$. For convergence it is necessary that f_{n+1} or (what amounts to the same thing) that f_n should tend to unity when n tends to infinity; it is therefore usual to write the typical factor f_n in the form $(1+u_n)$, where $u_n\to 0$ when $n\to\infty$. The form f_n is, however, useful, and $u_n=f_n-1$.

The condition that $u_n\to 0$ is necessary but *not sufficient* for convergence (§ 88, Ex. 1).

The two Lemmas that follow are often required.

Lemma 1. $1+a<e^a$ if $a>0$; $1-a<e^{-a}$ if $0<a<1$.

See § 25, Ex. 3.

Lemma 2. If $|u_n|=a_n$, whether u_n is real or complex,

$$|(1+u_{n+1})(1+u_{n+2})\ldots(1+u_{n+p})-1|$$
$$\leqq(1+a_{n+1})(1+a_{n+2})\ldots(1+a_{n+p})-1.$$

Take three factors,

$$1+u,\ 1+v,\ 1+w \quad\text{where}\quad |u|=a,\ |v|=b,\ |w|=c;$$

then $\quad (1+u)(1+v)(1+w)-1=(u+v+w)+(uv+uw+vw)+uvw,$

and therefore

$$|(1+u)(1+v)(1+w)-1|\leqq(a+b+c)+(ab+ac+bc)+abc,$$

that is $\quad \leqq(1+a)(1+b)(1+c)-1.$

The proof is obviously quite general.

88. Tests for Convergence of Products. Two tests for convergence will now be given; these will be sufficient for the applications we make, and Bromwich's treatise on *Infinite Series* may be consulted for further information.

Test 1. *If a_n is real and positive for every value of n the product $\Pi(1+a_n)$ converges or diverges according as the series Σa_n converges or diverges.*

Let $s_n=a_1+a_2+\ldots+a_n$ and, when Σa_n converges, let s_n tend to s when n tends to infinity.

(i) Let Σa_n be convergent. By Lemma 1 of § 87, $1+a_r<e^{a_r}$ and therefore

$$P_n=(1+a_1)(1+a_2)\ldots(1+a_n)<e^{a_1}e^{a_2}\ldots e^{a_n},$$

so that $\quad P_n<e^{s_n}<e^s.$

Thus P_n increases as n increases, but is less than the fixed number e^s for every value of n; therefore P_n tends to a limit and the product $\Pi(1+a_n)$ is convergent.

(ii) Let Σa_n be divergent. In this case

$$P_n=(1+a_1)(1+a_2)\ldots(1+a_n)=1+s_n+\text{(positive terms)},$$

and therefore $P_n\to+\infty$ when $n\to\infty$ since s_n does so. The infinite product $\Pi(1+a_n)$ is therefore divergent.

Test 2. *If u_n is any number, real or complex, and if $|u_n|=a_n$, the product $\Pi(1+u_n)$ converges if the product $\Pi(1+a_n)$ converges.*

By Lemma 2 of § 87

$$|(1+u_{n+1})(1+u_{n+2})\ldots(1+u_{n+p})-1|$$
$$\leqq(1+a_{n+1})(1+a_{n+2})\ldots(1+a_{n+p})-1.$$

Now (§ 87, Theorem) $[(1+a_{n+1})(1+a_{n+2})\ldots(1+a_{n+p})-1]$ tends to zero when n tends to ∞ if the product $\Pi(1+a_n)$ converges; therefore, when $\Pi(1+a_n)$ converges,

$$|(1+u_{n+1})(1+u_{n+2})\ldots(1+u_{n+p})-1|$$

also tends to zero when n tends to infinity so that the product $\Pi(1+u_n)$ is convergent.

Definition. If u_n is real or complex the product $\Pi(1+u_n)$ is said to converge absolutely when the product $\Pi(1+|u_n|)$ converges.

Hence $\Pi(1+u_n)$ converges absolutely if $\Sigma|u_n|$ converges and Πf_n converges absolutely if $\Sigma|f_n-1|$ converges.

Again, since $\Pi(1+u_n)$ cannot converge unless $u_n\to 0$ when $n\to\infty$, it may always be assumed in testing for convergence that $|u_n|<1$; the omission of all factors in which $|u_n|\geqq 1$ would at most affect the *value* of the product and not the property of convergence or divergence.

Suppose now that $\Pi(1+u_n)$ is expressed in the form

$$P\equiv\prod_1^{\infty}(1+u_n)=\prod_1^{m}(1+u_n)\times\prod_{m+1}^{\infty}(1+u_n)=P_m\,.\,Q_m$$

where
$$Q_m=\prod_{m+1}^{\infty}(1+u_n).$$

The product P will or will not converge according as the product Q_m does or does not converge. Further, we may suppose $|u_n|$ to be not merely less than unity but less than any positive number c when considering the convergence of

Q_m, the integer m being taken large enough to make $|u_n|<c$ when $n>m$.

An expression for log P will now be found and the assumption is expressly made that *every logarithm has its principal value.*

Suppose (i) that the series

$$\sum_{m+1}^{\infty} \log(1+u_n) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (\alpha)$$

converges, its sum being l; (ii) that m is so large that $|l|<\pi$; and, therefore, (iii) that l is the principal value of $\log Q_m$. It follows that

$$P = P_m \,.\, Q_m = (1+u_1)(1+u_2) \ldots (1+u_m)e^l$$

since $Q_m = e^l$.

Again, by § 70,

$$\log[(1+u_1)(1+u_2)\ldots(1+u_m)] = \sum_1^m \log(1+u_n) + 2k\pi i,$$

where k is zero or a positive or negative integer. Hence, since $l = \log Q_m$,

$$\log P = \sum^m \log(1+u_n) + 2k\pi i + l$$

$$= \sum^{\infty} \log(1+u_n) + 2k\pi i. \quad .(\beta)$$

The number k is not zero, in general, but it is finite (not greater than m numerically); even when u_n is real the factor $(1+u_n)$ may be negative and therefore $\log(1+u_n)$ may be complex for several values of n.

If there is no value of m for which the series (α) converges then the product Q_m and therefore also the product P cannot converge. The existence of the number l, that is, the convergence of the series (α), is therefore both sufficient and necessary for the convergence of the product $\Pi(1+u_n)$.

Note. In testing the convergence of $\Sigma|u_n|$ a useful comparison series is $\Sigma(1/n^2)$. Thus $\Sigma|u_n|$ converges if $n^2|u_n| \to k$, a constant, when $n \to \infty$; for if $k' > k$ and n sufficiently large $n^2|u_n|$ will be less than k' and therefore $|u_n| < k'/n^2$. But $\Sigma(1/n^2)$ converges and therefore $\Sigma|u_n|$ will also converge.

Ex. 1. (i) $\Pi\left(1 \pm \frac{1}{n^2}\right)$; (ii) $\Pi\left(1 \pm \frac{x^2}{n^2}\right)$; (iii) $\Pi\left(1 + \frac{1}{n}\right)$.

The products (i) and (ii) converge absolutely (x may be real or complex) because $\Sigma(1/n^2)$ converges. The product (iii) diverges because

$\Sigma(1/n)$ is a divergent series. The product (iii) is an example of a divergent product $\Pi(1+u_n)$ for which $u_n \to 0$ when $n \to \infty$ so that, as for infinite series, the condition that $u_n \to 0$ when $n \to \infty$ is necessary but not sufficient for convergence.

Ex. 2. If a_n is positive and less than unity for every value of n and Σa_n divergent the product $\Pi(1-a_n)$ diverges to zero.

Of course it would be the same thing if $a_n < 1$ for $n \geq m$, some fixed number. Now

$$1 - a_r = \frac{1-a_r^2}{1+a_r}, \text{ so that } 0 < 1 - a_r < \frac{1}{1+a_r} \text{ and therefore}$$

$$0 < \prod_{r=1}^{n} (1-a_r) < 1 \div \prod_{r=1}^{n} (1+a_r).$$

But the product $\prod_{1}^{n} (1+a_r)$ tends to $+\infty$ when $n \to \infty$ and therefore the product $\prod_{1}^{n} (1-a_r)$ tends to zero when $n \to \infty$.

Ex. 3. If $P_n = \dfrac{x(x+1)(x+2)\dots(x+n)}{y(y+1)(y+2)\dots(y+n)}$, where $0 < x < y$, show that $P_n \to 0$ when $n \to \infty$.

Here $\dfrac{x+n}{y+n} = 1 - \dfrac{y-x}{y+n}$. Let $\dfrac{y-x}{y+n} = a_n$, and the result follows from Ex. 2.

Ex. 4. If $\dfrac{a_n}{a_{n+1}} = 1 + \dfrac{b_n}{n}$ and if $b_n \to b > 0$ when $n \to \infty$, show that $a_n \to 0$ when $n \to \infty$.

$$\frac{a_1}{a_n} = \frac{a_1}{a_2} \cdot \frac{a_2}{a_3} \dots \frac{a_{n-2}}{a_{n-1}} \cdot \frac{a_{n-1}}{a_n} = \prod_{r=1}^{n-1} \left(1 + \frac{b_r}{r}\right).$$

When r is large b_r differs but little from the limit b, say $b_r > b' > 0$ when $r > m$. The series $\Sigma(b'/n)$ diverges to ∞ so that a_1/a_n diverges to ∞ and therefore $a_n \to 0$.

Ex. 5. The product $\Pi\left\{\left(1+\dfrac{x}{n}\right)e^{-\frac{x}{n}}\right\}$ converges absolutely for every value of x, real or complex.

By § 68 we have

$$\left(1+\frac{x}{n}\right)e^{-\frac{x}{n}} = \left(1+\frac{x}{n}\right)\left(1 - \frac{x}{n} + \tfrac{1}{2}\frac{x^2}{n^2} - \frac{1}{3!}\frac{x^3}{n^3} + \dots\right)$$

$$= 1 - \tfrac{1}{2}\frac{x^2}{n^2} + \tfrac{1}{3}\frac{x^3}{n^3} - \dots.$$

Hence
$$f_n - 1 = \left(1+\frac{x}{n}\right)e^{-\frac{x}{n}} - 1 = -\tfrac{1}{2}\frac{x^2}{n^2} + \tfrac{1}{3}\frac{x^3}{n^3} - \dots,$$

and therefore $n^2 |f_n - 1| \to \tfrac{1}{2}|x|^2$ when $n \to \infty$ so that the series $\Sigma |f_n - 1|$ converges and therefore the product Πf_n converges absolutely for every value of x, real or complex.

This example is of fundamental importance for what follows.

89. Derangement of Factors in Product. It is proved in § 59 that no derangement of the terms of an absolutely convergent series alters the sum of the series; it will now be proved that no derangement of the factors of an absolutely convergent product alters the value of the product.

Let the sequence $g_1, g_2, g_3, \ldots$ be a derangement of the sequence $f_1, f_2, f_3, \ldots$ in the sense that every element in one sequence occurs once, and only once, in the other.

Suppose that Πf_n converges absolutely; then the series $\Sigma(f_n - 1)$ converges absolutely and, by § 59, no derangement of its terms alters the sum of the series. Hence the series $\Sigma(g_n - 1)$, which is derived from the series $\Sigma(f_n - 1)$ by a derangement of its terms, is absolutely convergent (with the same sum) and therefore the product Πg_n is absolutely convergent. It has now to be proved that Πf_n and Πg_n are equal.

Let
$$P_m = \prod_{s=1}^{m} f_s \text{ and } Q_n = \prod_{r=1}^{n} g_r.$$

However large m may be, n may be chosen so that Q_n contains all the factors that occur in P_m; but whatever integer r may be there is one, and only one, integer s such that $g_r = f_s$ and therefore the quotient Q_n/P_m contains only factors $f_\alpha, f_\beta, \ldots, f_\lambda$ such that the integers $\alpha, \beta, \ldots, \lambda$ are each greater than m. Thus

$$\left|\frac{Q_n}{P_m} - 1\right| = |f_\alpha f_\beta \ldots f_\lambda - 1|$$

and, since Πf_n is absolutely convergent, $|f_\alpha f_\beta \ldots f_\lambda - 1|$ tends to zero when $m \to \infty$. But when $m \to \infty$ so does n and therefore Q_n/P_m tends to unity and P_m and Q_n tend to Πf_n and Πg_n respectively; but Πf_n, being an absolutely convergent product, is not zero so that $\Pi g_n = \Pi f_n$.

EXERCISES X.

1. Show that

$$\text{(i)} \prod_{n=2}^{\infty}\left\{1 - \frac{2}{n(n+1)}\right\} = \frac{1}{3}; \quad \text{(ii)} \prod_{n=2}^{\infty}\left(\frac{n^3-1}{n^3+1}\right) = \frac{2}{3}.$$

2. Prove
$$\mathop{\mathcal{L}} \frac{1.3.5\ldots(2n-1)}{2.4.6\ldots 2n} = 0.$$

3. If $0 < x < y$ prove that

$$\frac{1}{y} + \sum_{n=1}^{\infty} \frac{x(x+1)\dots(x+n-1)}{y(y+1)\dots(y+n-1)(y+n)} = \frac{1}{y-x}.$$

4. If the factors $(1+1/r)$ and $(1-1/r)$ are multiplied by $e^{-1/r}$ and $e^{1/r}$ respectively, prove that the products

$$\prod_{r=2}^{\infty} \left\{\left(1+\frac{1}{r}\right)e^{-\frac{1}{r}}\right\} \quad \text{and} \quad \prod_{r=2}^{\infty} \left\{\left(1-\frac{1}{r}\right)e^{\frac{1}{r}}\right\}$$

are convergent while the products

$$\prod_{r=2}^{n} \left(1+\frac{1}{r}\right) \quad \text{and} \quad \prod_{r=2}^{n} \left(1-\frac{1}{r}\right)$$

diverge, when $n \to \infty$, to infinity and to zero respectively.

5. Prove that the equation

$$1 - x + \frac{x(x-1)}{2!} - \frac{x(x-1)(x-2)}{3!} + \dots + (\quad 1)^n \frac{x(x-1)\dots(x-n+1)}{n!}$$
$$= (1-x)\left(1-\frac{x}{2}\right)\left(1-\frac{x}{3}\right)\dots\left(1-\frac{x}{n}\right)$$

holds for all values of x and investigate the relation, when $n \to \infty$, between the series and the product.

6. If $$P_n(x) = \prod_{r=1}^{n} (1 + xu_r)$$

show that $$P_n(x) = 1 + \sum_{r=1}^{n} \sigma_r x^r$$

where σ_r is the sum of the products of $u_1, u_2, \dots u_n$ taken r at a time, and deduce that if Σu_n is absolutely convergent

$$\prod_{n=1}^{\infty} (1 + xu_n) = 1 + \sum_{n=1}^{\infty} \sigma_n x^n,$$

the series being absolutely convergent for every value of x.

[If $S_n = |u_1| + |u_2| + \dots + |u_n|$ and if $S_n \to S$ when $n \to \infty$, then

$$|P_n(x)| \leqq e^{S_n |x|}, \quad \mathop{\mathcal{L}}_{n\to\infty} \prod_{r=1}^{n} (1 + |xu_r|) \leqq e^{S|x|};$$

also $$|\sigma_r| < S_n^r/r! \text{ if } r > 1.]$$

7. If $|x| < 1$, show that

$$(1+x)(1+x^2)(1+x^4)\dots(1+x^{2^n})\dots = 1/(1-x).$$

8. Prove that if $|q| < 1$ the four products, $n = 1, 2, \dots,$

$$q_0 = \Pi(1 - q^{2n}), \qquad q_1 = \Pi(1 + q^{2n}),$$
$$q_2 = \Pi(1 + q^{2n-1}), \quad q_3 = \Pi(1 - q^{2n-1})$$

are absolutely convergent. Further,

$$q_0 q_3 = \Pi(1 - q^n), \quad q_1 q_2 = \Pi(1 + q^n), \; q_1 q_2 q_3 = 1.$$

9. If $f(x)=\Pi(1+q^{2n-1}x)$, show that $(1+qx)f(q^2x)=f(x)$,

and if
$$f(x)=1+\sum_1^\infty A_n x^n,$$
show that
$$q+q^2A_1=A_1,\ q^{2n-1}A_{n-1}+q^{2n}A_n=A_n,$$
and
$$A_1=\frac{q}{1-q^2},\ A_n=\frac{q^{n^2}}{(1-q^2)(1-q^4)\dots(1-q^{2n})}.$$

10. If $F(x)=\Pi\{(1+q^{2n-1}x)(1+q^{2n-1}/x)\}$, show that $qxF(q^2x)$ is equal to $F(x)$; then prove that $F(x)$ may be represented by a series
$$F(x)=B_0+\sum_{n=1}^\infty B_n(x^n+1/x^n)$$
where $B_n=B_0q^{n^2}$, $B_0=1/q_0$, q_0 being the product in Ex. 8.

11. If $|a|>1$, prove that
$$\prod_{n=1}^\infty\left(1-\frac{x}{a^n}\right)=1+\sum_{n=1}^\infty\frac{(-1)^n x^n}{(a-1)(a^2-1)\dots(a^n-1)}.$$

12. If $|a|<1$ and $|x|<1$, prove that
$$\prod_{n=0}^\infty\frac{1}{1-a^n x}=1+\sum_{n=1}^\infty\frac{x^n}{(1-a)(1-a^2)\dots(1-a^n)}.$$

90. Uniform Convergence. If the factors f_n are functions of a real variable x, say $f_n=1+u_n(x)$, the question of uniform convergence arises. It is sufficient for our purposes to consider the case that corresponds to the convergence of series when the M-Test applies. It is assumed therefore that

(i) each function $u_n(x)$ is defined for the closed interval (a, b);

(ii) $|u_n(x)|<M_n$ for every n, where M_n is independent of x;

(iii) ΣM_n is convergent.

If $P(x)=\Pi\{1+u_n(x)\}$ the convergence will be uniform, and, further, if each function $u_n(x)$ is continuous for $a\leqq x\leqq b$ the product $P(x)$ will be a continuous function of x for that range.

That the product converges both absolutely and uniformly follows from the fact that $\Sigma|u_n(x)|<\Sigma M_n$ and that the convergence of $\Sigma|u_n(x)|$ is therefore independent of x.

Let
$$P_n(x)=\prod_{r=1}^n\{1+u_n(x)\},\quad Q_n=\prod_{r=1}^n(1+M_n);$$
then $|u_n(x)|<M_n$ and therefore, by Lemma 2 of § 87 and the conditions for convergence, m may be chosen (and then kept fixed) so that
$$|P_{m+p}(x)-P_m(x)|<Q_{m+p}-Q_m<\varepsilon'<\varepsilon\ \dots\dots\dots\dots(1)$$

for every x such that $a \leqq x \leqq b$ and for $p=1, 2, 3, \ldots$. Now let $p \to \infty$; therefore

$$|P(x)-P_m(x)| \leqq \varepsilon' < \varepsilon, \quad a \leqq x \leqq b. \qquad \ldots\ldots(2)$$

Suppose now that each function $u_n(x)$ is continuous. If x_1 is any number in (a, b) we can choose h so that

$$|P_m(x)-P_m(x_1)| < \varepsilon \quad \text{if} \quad |x-x_1| < h \qquad \ldots\ldots(3)$$

because $P_m(x)$ is the product of a finite number of continuous factors.

Hence $\quad P(x)-P(x_1)$

$$=[P(x)-P_m(x)]-[P(x_1)-P_m(x_1)]+[P_m(x)-P_m(x_1)],$$

and therefore, by (2) and (3),

$$|P(x)-P(x_1)| < 3\varepsilon \quad \text{if} \quad |x-x_1| < h,$$

so that $P(x)$ is continuous.

Differentiation. If $u_n(x)$ is continuous for a given range, say for $a \leqq x \leqq b$, the derivative of $\log P(x)$ is given by the equation,

$$\frac{P'(x)}{P(x)} = \sum_1^\infty \frac{u_n'(x)}{1+u_n(x)}$$

when the following conditions are satisfied:

(i) $|1+u_n(x)| > A > 0$, for $a \leqq x \leqq b$, and for every value of n;

(ii) $|u_n'(x)| < B_n$, independent of x, for $a \leqq x \leqq b$ and for every value of n;

(iii) ΣB_n convergent.

When these conditions are satisfied the series

$$\sum_1^\infty \frac{u_n'(x)}{1+u_n(x)}$$

converges uniformly for the range $a \leqq x \leqq b$ and therefore (*E.T.* p. 400) the series is the derivative of the series for $\log P(x)$.

The above conditions are not very wide, but they are sufficient for many applications.

Ex. If $\quad P(x)=x\prod_1^\infty\left(1+\frac{x^2}{n^2\pi^2}\right)$, show that

$$\frac{P'(x)}{P(x)}=\frac{1}{x}+\sum_1^\infty \frac{2x}{x^2+n^2\pi^2}, \quad 0 < a \leqq |x| \leqq b$$

where b is arbitrarily large.

$u_n(x)=x^2/n^2\pi^2$ and therefore $P(x)$ converges uniformly for every value of x, $|x| \leqq b$. If $v_n(x)=2x/(x^2+n^2\pi^2)$ the limit of $n^2v_n(x)$ for

$n \to \infty$ is $2x/\pi^2$; B_n may be taken to be $2b/n^2\pi^2$ and ΣB_n is convergent. Obviously $1+u_n(x)$ is positive for every value of n and x. The term $1/x$ requires that $|x|$ be positive.

91. Tannery's Theorem. The Theorem for products corresponding to that of § 63 for series may be stated as follows:

If $F(n) = \prod_{r=0}^{N} \{1+u_r(n)\}$, where N is a function of n that tends to infinity with n and $u_r(n)$ is a function of n, the product $F(n)$ will tend to a limit when n tends to infinity provided the following conditions are satisfied:

(i) $\underset{n\to\infty}{\mathcal{L}}\, u_r(n) = v_r$, when r is fixed;

(ii) $|u_r(n)| \leqq M_r$, where M_r is independent of n;

(iii) ΣM_r is convergent.

When these conditions are satisfied $F(n)$ tends to a limit when $n \to \infty$ and the limit is given by the equation

$$\underset{n\to\infty}{\mathcal{L}}\, F(n) = \prod_{r=0}^{\infty} (1+v_r).$$

As in § 63, it is plain that $\Sigma\,|v_r|$ converges and therefore the product $\Pi(1+v_r)$ is convergent. Again, since n and therefore also N is to tend to infinity, we may always suppose that N is greater than any given integer m, however large m may be. Now take the notations

$$P_s(n) = \prod_{r=0}^{s} \{1+u_r(n)\},\quad Q_s = \prod_{r=0}^{s} (1+v_r),\quad Q = \prod_{r=0}^{\infty} (1+v_r),$$

and express $F(n) - Q$ in the form $\alpha + \beta - \gamma$ where

$$\alpha = P_m(n) - Q_m,\ \beta = P_N(n) - P_m(n),\ \gamma = Q - Q_m.$$

We now have

$$|\beta| = |P_m(n)|\left|\prod_{r=m+1}^{N} \{1+u_r(n)\} - 1\right|$$

$$\leqq \prod_{r=0}^{m} (1+M_r)\left[\prod_{r=m+1}^{N} (1+M_r) - 1\right],$$

$$|\gamma| = |Q_m|\left|\prod_{r=m+1}^{\infty} (1+v_r) - 1\right| \leqq \prod_{r=0}^{m} (1+M_r)\left[\prod_{r=m+1}^{\infty} (1+M_r) - 1\right].$$

Since ΣM_r and therefore $\Pi(1+M_r)$ converges m may be chosen so that, given ε as usual, both $|\beta|$ and $|\gamma|$ will be less

than ε. When m has been chosen let it be kept fixed and then n_1 may be chosen so that if $n > n_1$ the modulus

$$|\alpha| = |P_m(n) - Q_m|$$

will, by condition (i), be less than ε. Hence

$$|F(n) - Q| < 3\varepsilon \text{ if } n > n_1, \quad \text{or} \quad \mathop{\mathcal{L}}_{n\to\infty} F(n) = \prod_0^\infty (1 + v_r).$$

This Theorem is essentially the same as that stated by Chrystal on p. 346, Part II, of his *Algebra*.

92. Infinite Products for Trigonometric Functions. The expression of $\sin x$, $\sinh x$ and similar functions as infinite products was given by Euler in his *Analysis Infinitorum*, Vol. I, §§ 156 *et seq.*; the following method, which is an improved version of Euler's, is given by Tannery and Molk, *Fonctions Elliptiques*, I, Chapter III of the *Introduction*, and is said to be due to Darboux.

Let
$$f_n(x) = \tfrac{1}{2}\left\{\left(1 + \frac{x}{n}\right)^n - \left(1 - \frac{x}{n}\right)^n\right\} \qquad \text{.................(1)}$$

where n is an *odd* positive integer and x is any number, real or complex; $f_n(x) \to \sinh x$ when $n \to \infty$.

$f_n(x)$ is a polynomial of degree n in x; the absolute term of the polynomial is zero and the coefficient of the first power of x is unity so that $f_n(x)/x \to 1$ when $x \to 0$.

The roots of $f_n(x) = 0$ are 0 and $\pm x_k$ where

$$x_k = in\tan(k\pi/n),\ k = 1, 2, \dots, \tfrac{1}{2}(n-1) = N,$$

and therefore

$$f_n(x) = Ax\prod_{k=1}^{N}\left\{x^2 + n^2\tan^2\left(\frac{k\pi}{n}\right)\right\},\quad A = \text{constant}.$$

But $f_n(x)/x \to 1$ when $x \to 0$; therefore

$$1 = A\prod_1^N\left\{n^2\tan^2\left(\frac{k\pi}{n}\right)\right\}$$

and
$$f_n(x) = x\prod_{k=1}^{N}\left\{1 + \frac{x^2}{n^2\tan^2(k\pi/n)}\right\},\quad N = \tfrac{1}{2}(n-1) \qquad (2)$$

We now apply Tannery's Theorem, § 91. The greatest value of $k\pi/n$ is $(n-1)\pi/2n$ which is less than $\pi/2$ so that $n^2\tan^2(k\pi/n)$ is greater than $k^2\pi^2$ and therefore, if $|x^2| = a^2$

$$|\{1 + x^2/n^2\tan^2(k\pi/n)\}| < 1 + a^2/k^2\pi^2,$$

and the series $\Sigma a^2/k^2\pi^2$ converges. Further, when k is fixed,

$$\lim_{n\to\infty} \{1 + x^2/n^2 \tan^2(k\pi/n)\} = 1 + x^2/k^2\pi^2.$$

The conditions of Tannery's Theorem are therefore satisfied; and since $f_n(x) \to \sinh x$ when $n \to \infty$, we find

$$\sinh x = x\prod_{k=1}^{\infty}(1 + x^2/k^2\pi^2) = x\prod_{1}^{\infty}(1 + x^2/n^2\pi^2), \quad \text{.........(3)}$$

and the product converges absolutely and uniformly for the range $|x| \leqq K$ where K is an arbitrarily large positive number. The number M_n of § 90 may be taken to be $K^2/n^2\pi^2$.

Again, $\cosh x = \sinh 2x/2\sinh x$. The product (3) is absolutely convergent, and therefore the factors of the product may be arranged so that one set contains x and the even multiples of π while the other contains the odd multiples of π; thus

$$\sinh 2x = 2x\prod_{n=1}^{\infty}\left(1 + \frac{4x^2}{4n^2\pi^2}\right) \times \prod_{n=1}^{\infty}\left\{1 + \frac{4x^2}{(2n-1)^2\pi^2}\right\},$$

and therefore

$$\cosh x = \prod_{n=1}^{\infty}\left\{1 + \frac{4x^2}{(2n-1)^2\pi^2}\right\}. \quad \text{.........................(4)}$$

The formulae (3) and (4) are valid for complex as well as for real values of x and therefore if ix is substituted for x we find

$$\sin x = x\prod_{n=1}^{\infty}\left(1 - \frac{x^2}{n^2\pi^2}\right) \quad \text{.......................(5)}$$

$$\cos x = \prod_{n=1}^{\infty}\left\{1 - \frac{4x^2}{(2n-1)^2\pi^2}\right\}. \quad \text{..............(6)}$$

93. Expansions in Partial Fractions. The products in (3), (4), (5) and (6) of the preceding article may be differentiated logarithmically (§ 90). Thus from (5), if $0 < x < \pi$,

$$\cot x = \frac{1}{x} + \sum_{n=1}^{\infty}\frac{2x}{x^2 - n^2\pi^2} \quad \text{.....................(1)}$$

$$= \sum_{n=0}^{\infty}\left\{\frac{1}{n\pi + x} - \frac{1}{(n+1)\pi - x}\right\}. \quad \text{......(1a)}$$

Again, $\cot \frac{1}{2}x - \cot x = 1/\sin x$; therefore

$$\frac{1}{\sin x} = \frac{1}{x} + \sum_{n=1}^{\infty}(-1)^n\frac{2x}{x^2 - n^2\pi^2} \quad \text{.......................(2)}$$

$$= \sum_{n=0}^{\infty}(-1)^n\left\{\frac{1}{n\pi + x} + \frac{1}{(n+1)\pi - x}\right\}. \quad \text{......(2a)}$$

Another method of obtaining an expansion in partial fractions is given by Tannery and Molk (see § 92) and is, like the method for obtaining an infinite product, said to be due to Darboux. The method may be seen by taking the function $1/\sinh x$, defined as the limit of $1/f_n(x)$, where $f_n(x)$ is the polynomial of § 92.

Express $1/f_n(x)$ as a sum of partial fractions; since $f_n(x)/x \to 1$ when $x \to 0$ we have

$$\frac{1}{f_n(x)}=\frac{1}{x}+\sum_{k=1}^{N}\frac{A_k}{x-x_k}+\sum_{k=1}^{N}\frac{B_k}{x+x_k}, \quad N=\tfrac{1}{2}(n-1).$$

Now (*E.T.* p. 291) if $f_n'(x)=df_n(x)/dx$ the value of A_k is $1/f_n'(x_k)$; therefore, as is easily proved,

$$A_k=(-1)^k\left(\cos\frac{k\pi}{n}\right)^{n-2}=B_k,$$

so that
$$\frac{1}{f_n(x)}=\frac{1}{x}+\sum_{k=1}^{N}\frac{2xA_k}{x^2-x_k^2}. \quad \text{.............................(3)}$$

Apply Tannery's Theorem, § 63. If k is fixed

$$\mathop{\mathcal{L}}_{n\to\infty}\frac{2xA_k}{x^2-x_k^2}=(-1)^k\frac{2x}{x^2+k^2\pi^2}$$

(For the limit of A_k see § 25, Ex. 5.)

Next $|A_k|<1$ and $|x^2-x_k^2|>|k^2\pi^2-|x|^2|$ so that

$$\left|\frac{2xA_k}{x^2-x_k^2}\right|<\frac{2|x|}{|k^2\pi^2-|x|^2|}.$$

If X is not a multiple of π the series

$$\sum_{k=1}^{\infty}\frac{2X}{|k^2\pi^2-X^2|}$$

is absolutely convergent and therefore the conditions required by Tannery's Theorem are satisfied, and we find

$$\frac{1}{\sinh x}=\frac{1}{x}+\sum_{n=1}^{\infty}(-1)^n\frac{2x}{x^2+n^2\pi^2}. \quad \text{...............(4)}$$

The values 0, $\pm n\pi i$ are of course not values that x may take.

In (4) put ix in place of x and the series (2) for $1/\sin x$ is obtained.

By either of the above methods various series may be derived; as the product formulae hold whether x be real or complex the series for a trigonometric function can at once be transformed into one for the corresponding hyperbolic function

and *vice versa*. An identity such as $\tan x = \cot x - 2\cot 2x$ is also useful. Several of these series are given in Exercises XI and many developments which lie outside our limits will be found in Chrystal's *Algebra*, Chapter XXX, Bromwich's *Infinite Series*, and books on Trigonometry, such as Hobson's.

Ex. 1. Show that if $0 < x < \pi$

$$\frac{1}{\sin x} = \sum_{r=0}^{n}(-1)^r\left(\frac{1}{r\pi + x} + \frac{1}{(r+1)\pi - x}\right) + (-1)^{n+1}R_n(x)$$

where $$R_n(x) < 2/(n+1)\pi.$$

Write equation (2a) in the form

$$\frac{1}{\sin x} = \sum_{r=0}^{n}(-1)^r \quad \frac{1}{r\pi + x} + \frac{1}{(r+1)\pi - x} \quad + (-1)^{n+1}R_n(x),$$

where $R_n(x) = u_{n+1} - u_{n+2} + \dots,\ u_r = \dfrac{1}{r\pi + x} + \dfrac{1}{(r+1)\pi - x}.$

If $0 \leqq x \leqq \pi$, $u_r > u_{r+1}$, $r = n+1, n+2, \dots$ and therefore

$$R_n(x) < u_{n+1} < 2/(n+1)\pi.$$

Ex. 2. Show that

$$\cos x - \cos\alpha = (1-\cos\alpha)\left(1 - \frac{x^2}{\alpha^2}\right)\prod_1^\infty\left\{\left(1 - \frac{x^2}{(2n\pi - \alpha)^2}\right)\left(1 - \frac{x^2}{(2n\pi + \alpha)^2}\right)\right\}$$

$$(\cos x - \cos\alpha) = 2\sin\frac{\alpha - x}{2}\sin\frac{\alpha + x}{2}$$

$$= \frac{\alpha^2 - x^2}{2}\prod_1^\infty\left\{\left(1 - \frac{(\alpha - x)^2}{4n^2\pi^2}\right)\left(1 - \frac{(\alpha + x)^2}{4n^2\pi^2}\right)\right\}.$$

Let $x = 0$; $1 - \cos\alpha = \dfrac{\alpha^2}{2}\Pi\left\{\left(1 - \dfrac{\alpha^2}{4n^2\pi^2}\right)\left(1 - \dfrac{\alpha^2}{4n^2\pi^2}\right)\right\}.$

Take the quotient $(\cos x - \cos\alpha)/(1 - \cos\alpha)$; the typical factor is

$$\frac{4n^2\pi^2 - (\alpha - x)^2}{4n^2\pi^2 - \alpha^2}\cdot\frac{4n^2\pi^2 - (\alpha + x)^2}{4n^2\pi^2 - \alpha^2}$$

$$= \left(1 + \frac{x}{2n\pi - \alpha}\right)\left(1 - \frac{x}{2n\pi + \alpha}\right)\left(1 - \frac{x}{2n\pi - \alpha}\right)\left(1 + \frac{x}{2n\pi + \alpha}\right)$$

$$= \left(1 - \frac{x^2}{(2n\pi - \alpha)^2}\right)\left(1 - \frac{x^2}{(2n\pi + \alpha)^2}\right).$$

Ex. 3. Find the values of $\displaystyle\sum_{n=1}^\infty \frac{1}{n^2},\ \sum_{n=1}^\infty \frac{1}{n^4}.$

From the infinite product for $\sin x/x$ we find

$$-\log\frac{\sin x}{x} = -\sum_{m=1}^\infty \log\left(1 - \frac{x^2}{m^2\pi^2}\right).$$

Now if $x^2 < \pi^2$ and A_m is the series

$$A_m = -\log\left(1 - \frac{x^2}{m^2\pi^2}\right) = \frac{x^2}{m^2\pi^2} + \frac{1}{2}\frac{x^4}{m^4\pi^4} + \frac{1}{3}\frac{x^6}{m^6\pi^6} + \dots.$$

the series A_m and ΣA_m satisfy the conditions for derangement of § 66. Hence we have

$$-\sum_{m=1}^{\infty} \log\left(1-\frac{x^2}{m^2\pi^2}\right)=c_2x^2+c_4x^4+c_6x^6+\dots \qquad \text{.............(1)}$$

where
$$c_2=\frac{1}{\pi^2}\sum_1^{\infty}\frac{1}{n^2},\ c_4=\frac{1}{2\pi^4}\sum_1^{\infty}\frac{1}{n^4}\dots$$

Again, since $\sin x/x=1-x^2/6+x^4/5!-$ etc.

$$-\log\left(\frac{\sin x}{x}\right)=-\log\left\{1-\left(\frac{x^2}{6}-\frac{x^4}{120}+\dots\right)\right\}$$
$$=\left(\frac{x^2}{6}-\frac{x^4}{120}+\dots\right)+\tfrac{1}{2}\left(\frac{x^2}{6}-\frac{x^4}{120}+\dots\right)^2+,$$

and, as before the series may be re-arranged in powers of x, so that

$$-\log\left(\frac{\sin x}{x}\right)=\frac{x^2}{6}+\frac{x^4}{180}+\dots \qquad \text{.......................(2)}$$

The series (1) and (2) are convergent for a common range, say for $x^2<a^2$; therefore equating coefficients we find

$$c_2=\frac{1}{6}\quad\text{or}\quad\sum_1^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6},$$
$$c_4=\frac{1}{180}\quad\text{or}\quad\sum_1^{\infty}\frac{1}{n^4}=\frac{\pi^4}{90}.$$

In the same way the values of Σn^{-6}, Σn^{-8}, ... may be found.

Ex. 4. Show that $\Sigma(2n-1)^{-2}=\pi^2/8$, $\Sigma(2n-1)^{-4}=\pi^4/96, \dots$.
Proceed as in Ex. 3, using the infinite product for $\cos x$.

94. Bernoulli's Numbers. In equation (1) of § 93 put ix for x; then we find

$$\coth x=\frac{1}{x}+\sum_{n=1}^{\infty}\frac{2x}{n^2\pi^2+x^2}, \qquad \text{..............(1)}$$

and if we now put $\frac{1}{2}t$ for x the equation (1) gives, after a slight reduction,

$$\frac{t}{e^t-1}=1-\tfrac{1}{2}t+2\sum_{n=1}^{\infty}\frac{t^2}{4n^2\pi^2+t^2}. \qquad \text{.................(2)}$$

The series in (2) may be expressed as a series in powers of t; for if $|t|=\alpha<2\pi$ we have, $n=1, 2, 3, \dots$,

$$\frac{t^2}{4n^2\pi^2+t^2}=\frac{t^2}{4n^2\pi^2}\left(1+\frac{t^2}{4n^2\pi^2}\right)^{-1}$$
$$=\frac{\quad}{4n^2\pi^2}-\frac{\quad}{(4n^2\pi^2)^2}+\dots+(-1)^{m-1}\cdot\frac{t^{2n}}{(4n^2\pi^2)^m}+\dots.$$

This series converges absolutely if $|t| = \alpha < 2\pi$, and therefore, by § 66, the series in (2), namely,

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(-1)^{m-1}\frac{t^{2m}}{(4n^2\pi^2)^m}$$

may be evaluated by summing first with respect to n and next with respect to m. Thus

$$\sum_{n=1}^{\infty}\frac{t^2}{4n^2\pi^2+t^2} = \sum_{m=1}^{\infty}(-1)^{m-1}\frac{t^{2m}}{(2\pi)^{2m}}\sum_{n=1}^{\infty}\frac{1}{n^{2m}}$$

$$= \sum_{m=1}^{\infty}(-1)^{m-1}\frac{S_{2m}}{(2\pi)^{2m}}t^{2m}$$

if
$$S_{2m} = \frac{1}{1^{2m}} + \frac{1}{2^{2m}} + \ldots = \sum_{n=1}^{\infty}\frac{1}{n^{2m}}. \quad \text{..........(3)}$$

Hence, if we now, for convenience, interchange n and m we get

$$\frac{t}{e^t-1} = 1 - \tfrac{1}{2}t + 2\sum_{n=1}^{\infty}(-1)^{n-1}\frac{S_{2n}}{(2\pi)^{2n}}t^{2n}. \quad \text{..........(4)}$$

Again (*E.T.* p. 404, Ex. 7)

$$\frac{t}{e^t-1} = 1 - \tfrac{1}{2}t + \sum_{n=1}^{\infty}(-1)^{n-1}\frac{B_n}{(2n)!}t^{2n} \quad \text{..........(5)}$$

where B_1, B_2, ... are Bernoulli's numbers. Therefore, equating the coefficients of t^{2n} in (4) and (5) we find

$$B_n = 2\frac{(2n)!}{(2\pi)^{2n}}S_{2n}. \quad \text{..........(6)}$$

It may be verified from this expression for B_n that the series (5) converges absolutely when $|t| < 2\pi$; for $S_{2n+2} < S_{2n}$ and

$$\left| -\frac{B_{n+1}}{B_n}\cdot\frac{t^2}{(2n+1)(2n+2)}\right| = \frac{S_{2n+2}}{S_{2n}}\cdot\left|\frac{t}{2\pi}\right|^2 < \left|\frac{t}{2\pi}\right|^2.$$

Equation (4) shows that $t/(e^t-1)$ is expressible by a power series and therefore the Maclaurin series (5) is now justified; also (5) holds for complex values of t since in (4) t may be complex.

The remainder after the term in t^{2n}, equation (5), may be put in a convenient form, t being *real*. For

$$\frac{2t^2}{4n^2\pi^2+t^2} = 2\sum_{r=1}^{m}(-1)^{r-1}\cdot\frac{t^{2r}}{(4n^2\pi^2)^r} + (-1)^m\cdot\frac{2t^{2m+2}}{(4n^2\pi^2+t^2)(4n^2\pi^2)^m},$$

and therefore, when summation is made with respect to n in the series (2), the coefficient of $(-1)^m t^{2m+2}$ is

$$2\sum_{n=1}^{\infty}\frac{1}{(4n^2\pi^2+t^2)(4n^2\pi^2)^m}<2\sum_{n=1}^{\infty}\frac{1}{(4n^2\pi^2)^{m+1}}=\frac{2S_{2m+2}}{(2\pi)^{2m+2}},$$

so that the coefficient is of the form $\theta_m\{2S_{2m+2}/(2\pi)^{2m+2}\}$ where $0<\theta_m<1$.

Hence, interchanging m and n as before, we see that the remainder $R_n(t)$ after the term in t^{2n} in (5) is given by

$$R_n(t)=(-1)^n\theta_n\frac{B_{n+1}}{(2n+2)!}t^{2n+2},\ 0<\theta_n<1. \quad\text{..........(5a)}$$

If t is complex the above reasoning fails, but there is a similar form.

Since
$$\frac{t}{e^t+1}=\frac{t}{e^t-1}-\frac{2t}{e^{2t}-1},$$

we deduce from (5) that if $|t|<\pi$,

$$\frac{t}{e^t+1}=\tfrac{1}{2}t+\sum_{n=1}^{\infty}(-1)^n\frac{2^{2n}-1}{(2n)!}B_n t^{2n},\ |t|<\pi, \quad\text{.....(7)}$$

and, by putting ix for t, series expressing various trigonometric functions in terms of $B_1, B_2, \ldots$ may be found (see Exercises XI).

The following values of B_n may be useful :

$$B_1=\tfrac{1}{6},\ B_2=\tfrac{1}{30},\ B_3=\tfrac{1}{42},\ B_4=\tfrac{1}{30},\ B_5=\tfrac{5}{66},\ B_6=\tfrac{691}{2730}.$$

See Chrystal's *Algebra*, Part II, Chapter XXVIII, § 6; Nielsen, *Traité Élémentaire des Nombres de Bernoulli.*

For the expression of B_n as an integral see § 165, Ex. 5.

EXERCISES XI.

1. Show that $\prod_{n=1}^{\infty}\left(\frac{n^2+1}{n^2}\right)=(e^{\pi}-e^{-\pi})/2\pi.$

2. Show that $\prod_{n=1}^{\infty}\left(\frac{4n^2-4n+2}{4n^2-4n+1}\right)=\frac{1}{2}(e^{\frac{\pi}{2}}+e^{-\frac{\pi}{2}}).$

3. $1+\sin x=\frac{1}{8}(\pi+2x)^2\left\{1-\frac{(\pi+2x)^2}{4^2\pi^2}\right\}^2\left\{1-\frac{(\pi+2x)^2}{8^2\pi^2}\right\}^2\ldots$

4. (i) $\cos\frac{x}{2}\cos\frac{x}{2^2}\cos\frac{x}{2^3}\cos\frac{x}{2^4}\ldots=\frac{\sin x}{x};$

(ii) $\frac{1}{2}\tan\frac{x}{2}+\frac{1}{2^2}\tan\frac{x}{2^2}+\frac{1}{2^3}\tan\frac{x}{2^3}+\ldots=\frac{1}{x}-\cot x.$

What restriction is there on x in case (ii) ?

5. (i) $\dfrac{\sin x}{x}=\prod_{n=1}^{\infty}\left(1-\frac{4}{3}\sin^2\frac{x}{3^n}\right);$

(ii) $\cos x=\prod_{n=1}^{\infty}\left(1-4\sin^2\frac{x}{3^n}\right).$ (Laisant)

6. $\dfrac{\sin 3x}{3\sin x}=\prod_{n=1}^{\infty}\left\{\left(1+\frac{2x}{n\pi+x}\right)\left(1-\frac{2x}{n\pi-x}\right)\right\}.$

Examples 7-17 are taken from Chapter IX of Euler's *Introductio in Analysin Infinitorum.*

7. $\dfrac{e^{b+x}+e^{c-x}}{e^b+e^c}=\prod_{n=1}^{\infty}\left\{1+\frac{4(b-c)x+4x^2}{(2n-1)^2\pi^2+(b-c)^2}\right\}.$

8. $\dfrac{e^{b+x}-e^{c-x}}{e^b-e^c}=\left(1+\frac{2x}{b-c}\right)\prod_{n=1}^{\infty}\left\{1+\frac{4(b-c)x+4x^2}{4n^2\pi^2+(b-c)^2}\right\}$

9. $(\cosh x+\cosh c)/(1+\cosh c)$

$$=\prod_{n=1}^{\infty}\left\{\left(1+\frac{2cx+x^2}{(2n-1)^2\pi^2+c^2}\right)\left(1-\frac{2cx-x^2}{(2n-1)^2\pi^2+c^2}\right)\right\}.$$

10. $(\cosh x-\cosh c)/(1-\cosh c)$

$$=\left(1-\frac{x^2}{c^2}\right)\prod_{n=1}^{\infty}\left\{\left(1+\quad 2cx+x^2\quad\right)\left(1-\frac{2cx-x^2}{4n^2\pi^2+c^2}\right)\right\}.$$

11. $(\sinh x+\sinh c)/\sinh c$

$$=\left(1+\frac{x}{c}\right)\prod_{n=1}^{\infty}\left\{1+\frac{(-1)^n 2cx+x^2}{n^2\pi^2+c^2}\right\}.$$

12. $(\sinh x-\sinh c)/\sinh c$

$$=-\left(1-\frac{x}{c}\right)\prod_{n=1}^{\infty}\left\{1+\frac{(-1)^{n-1}2cx+x^2}{n^2\pi^2+c^2}\right\}.$$

13. By putting ix for x and ic for c, or otherwise, deduce the formulae for the circular functions corresponding to those of Examples 9-12. For instance, from 9, if $m=2n-1$

$$\frac{\cos x+\cos c}{1+\cos c}=\prod_{n=1}^{\infty}\left\{\left(1-\frac{x^2}{(m\pi+c)^2}\right)\left(1-\frac{x^2}{(m\pi-c)^2}\right)\right\}.$$

$$\frac{\cos(x-c)}{\cos c}=\prod_{n=1}^{\infty}\left\{\left(1+\frac{2x}{(2n-1)\pi-2c}\right)\left(1-\frac{2x}{(2n-1)\pi+2c}\right)\right\}.$$

15. $\dfrac{\sin(c-x)}{\sin c}=\left(1-\frac{x}{c}\right)\prod_{n=1}^{\infty}\left\{\left(1+\frac{x}{n\pi-c}\right)\left(1-\frac{x}{n\pi+c}\right)\right\}.$

16. $\dfrac{\cosh x-\cos c}{1-\cos c}=\left(1+\frac{x^2}{c^2}\right)\prod_{n=1}^{\infty}\left\{\left(1+\frac{x^2}{(2n\pi-c)^2}\right)\left(1+\frac{x^2}{(2n\pi+c)^2}\right)\right\}.$

17. $\dfrac{\cosh x+\cos c}{1+\cos c}=\prod_{n=1}^{\infty}\left\{\left(1+\frac{x^2}{(\overline{2n-1}\pi+c)^2}\right)\left(1+\frac{x^2}{(\overline{2n-1}\pi-c)^2}\right)\right\}.$

18. From Examples 16, 17, deduce products for $\cosh 2u \pm \cos 2u$ and deduce the value of

$$\prod_{n=1}^{\infty}\left(\frac{4n^2-4n+5}{4n^2-4n+1}\right).$$

19. $\prod_{n=1}^{\infty}\left(1+\frac{x^4}{n^4}\right)=\frac{\cosh(\pi x\sqrt{2})-\cos(\pi x\sqrt{2})}{2\pi^2x^2}$

20. $\sin \pi x=\pi x(x+1)\prod_{n=1}^{\infty}\left\{1-\frac{x(x+1)}{n(n+1)}\right\}.$

21. $\coth x=\frac{1}{x}+\sum_{n=1}^{\infty}\frac{2x}{n^2\pi^2+x^2}.$

22. (i) $\frac{\cosh ax}{\sinh a\pi}=\frac{1}{a\pi}+\frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^n\frac{a\cos nx}{n^2+a^2};\quad -\pi\leqq x\leqq \pi;$

(ii) $\frac{\sinh ax}{\sinh a\pi}=\frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^{n-1}\frac{n\sin nx}{n^2+a^2};\quad -\pi<x<\pi;$

(iii) $\frac{\sinh x}{\cosh x+\cos c}=\sum_{n=1}^{\infty}\left\{\frac{2x}{[(2n-1)\pi-c]^2+x^2}+\frac{2x}{[(2n-1)\pi+c]^2+x^2}\right\}.$

23. $\tan x=\sum_{n=1}^{\infty}\frac{8x}{(2n-1)^2\pi^2-4x^2}.$

24. $\tanh x=\sum_{n=1}^{\infty}\frac{8x}{(2n-1)^2\pi^2+4x^2}.$

25. $\frac{1}{\cos x}=\sum_{n=1}^{\infty}(-1)^{n-1}\frac{4(2n-1)\pi}{(2n-1)^2\pi^2-4x^2}.$

26. $\frac{1}{\cosh x}=\sum_{n=1}^{\infty}(-1)^{n-1}\frac{4(2n-1)\pi}{(2n-1)^2\pi^2+4x^2}.$

27. $\frac{1}{a}+\sum_{n=1}^{\infty}\frac{2a}{a^2+n^2}=\pi\coth a\pi.$

Put $a\pi$ for x in Ex. 21. Many numerical series may be expressed in finite form by assigning particular values to x in Examples 21-26 and similar examples.

28. $x\cot x=1-\sum_{n=1}^{\infty}\frac{2^{2n}B_n}{(2n)!}x^{2n},\quad |x|<\pi.$

29. $\tan x=\sum_{n=1}^{\infty}(2^{2n}-1)\frac{2^{2n}B_n}{(2n)!}x^{2n-1},\quad |x|<\frac{\pi}{2}.$

30. $\frac{x}{\sin x}=1+\sum_{n=1}^{\infty}\frac{(2^{2n}-2)B_n}{(2n)!}x^{2n},\quad |x|<\pi.$

31. Derive the series for the hyperbolic functions corresponding to the circular functions in 28-30 by putting ix for x.

32. By multiplying the power series for $(e^t-1)/t$ and the series (5) of § 94 for $t/(e^t-1)$ prove that

$$\binom{2n+1}{2}B_1-\binom{2n+1}{4}B_2+\binom{2n+1}{6}B_3-\dots$$
$$+(-1)^{n-1}\binom{2n+1}{1}B_n=\frac{2n-1}{2}.$$

33. Show from equation (2) of § 94 that if $0<t$

$$0<\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}<\frac{t}{12}.$$

Note that $$\sum^{\infty}(4n^2\pi^2+t^2)^{-1}<\frac{1}{4\pi^2}\sum\frac{1}{n^2}.$$

34. Expand the logarithm of the infinite product (i) for $(\sin x/x)$ and (ii) for $\cos x$ as a double series, and show that it may in each case be arranged as a power series in x (see § 93, Ex. 3); then show that

(i) $$\log\frac{\sin x}{x}=-\sum_{n=1}^{\infty}\frac{2^{2n-1}B_n}{(2n)!\,n}x^{2n},\quad |x|<\pi;$$

(ii) $$\log\cos x=-\sum_{n=1}^{\infty}\frac{2^{2n-1}(2^{2n}-1)B_n}{(2n)!\,n}x^{2n},\quad |x|<\frac{\pi}{2}.$$

35. $$\sec x=\sum_{n=0}^{\infty}\frac{2^{2n+2}T_{2n+1}}{\pi^{2n+1}}x^{2n},\quad |x|<\frac{\pi}{2}.$$

$$=1+\sum_{n=1}^{\infty}\frac{E_n}{(2n)!},$$

where $$T_{2n+1}=\frac{1}{1^{2n+1}}-\frac{1}{3^{2n+1}}+\frac{1}{5^{2n+1}}-\frac{1}{7^{2n+1}}+\dots$$

and $$E_n=\frac{2^{2n+2}(2n)!\,T_{2n+1}}{\pi^{2n+1}}.$$

The numbers E_n are called Euler's Numbers. $E_1=1$, $E_2=5$, $E_3=61$, $E_4=1385$. (Chrystal's *Algebra*, Part II, Ch. XXX, § 3 and § 15.)

36. Prove that $1+\frac{1}{2m-1}>S_{2m}>1+\frac{1}{(2m-1)2^{2m-1}}$.

Apply the inequalities (5) of § 11.

37. Prove that B_n and E_n are both positive and tend to ∞ when $n\to\infty$.

Examples 38-40 are from Tannery and Molk, *Fonctions Elliptiques* I.

38. If λ is real and positive but not greater than unity, show that

$$\frac{\cos\lambda x}{\sin x}=\frac{1}{x}+\sum_{n=1}^{\infty}(-1)^n\frac{2x\cos n\lambda\pi}{x^2-n^2\pi^2}$$

$$=\mathop{\mathcal{L}}_{m\to\infty}\sum_{n=-m}^{m}(-1)^n\frac{\cos n\lambda\pi}{x-n\pi}.$$

Let $f_n(x) = i\varphi_n(x)/\psi_n(x)$ where n is an odd positive integer and $\varphi_n(x)$, $\psi_n(x)$ are polynomials in x, namely :

$$\varphi_n(x) = \left(1 + \frac{\lambda xi}{n}\right)^n + \left(1 - \frac{\lambda xi}{n}\right)^n, \quad \psi_n(x) = \left(1 + \frac{xi}{n}\right)^n - \left(1 - \frac{xi}{n}\right)^n;$$

$f_n(x) \to \cos \lambda x/\sin x$ when $n \to \infty$ and $xf_n(x) \to 1$ when $x \to 0$.

Express $f_n(x)$ in partial fractions ; the roots of $\psi_n(x) = 0$ are $x = 0$ and $x = \pm x_k = \pm n \tan (k\pi/n)$, $k = 1, 2, \ldots, \frac{1}{2}(n-1) = N$.

$$f_n(x) = \frac{1}{x} + \sum_{k=1}^{N} \left(\frac{A_k}{x - x_k} + \frac{B_k}{x + x_k}\right)$$

and $$A_k = B_k = (-1)^k \left[\frac{\cos (k\pi/n)}{\cos \alpha_k}\right]^n \cdot \frac{\cos (n\alpha_k)}{\cos^2(k\pi/n)}, \quad \tan \alpha_k = \lambda \tan \frac{k\pi}{n}.$$

Now proceed as in § 93. For the limit of A_k see Exercises II, 27. Show that the series represents the function if $-1 \leqq \lambda \leqq 1$.

39. $$\frac{\sin \lambda x}{\sin x} = \sum_{n=1}^{\infty} (-1)^n \frac{2n\pi \sin n\lambda\pi}{x^2 - n^2\pi^2} = \mathcal{L}_{\infty} \sum_{n=-m} (-1)^n \frac{\sin n\lambda\pi}{x - n\pi},$$

where λ is the same as in Ex. 38.

Note. $$\frac{\sin \lambda x}{\sin x} = \frac{1}{\sin \lambda\pi}\left\{\cos \lambda\pi \frac{\cos \lambda x}{\sin x} + \frac{\cos \lambda(x+\pi)}{\sin (x+\pi)}\right\}.$$

40. $$\frac{e^{\lambda xi}}{\sin x} = \mathcal{L}_{\infty} \sum_{n=-m} (-1)^n \; e^{n\lambda\pi}$$

95. The Gamma Function.

The product $P_n(x)$ where

$$P_n(x) = \frac{n!\, n^{x-1}}{x(x+1)(x+2)\ldots(x+n-1)} \quad \ldots\ldots\ldots \quad (1)$$

is defined for all values of x, real or complex, *except* the values zero, and the negative integers numerically less than n. It will now be shown that $P_n(x)$ tends to a limit when n tends to infinity ; the limit is called the *Gamma Function* of x, is denoted by $\Gamma(x)$ and is defined for all values of x, real or complex, except zero and the negative integers.

The limit will not be altered if $P_n(x)$ is multiplied by the factor $n/(n+x)$ which tends to unity when n tends to infinity; hence we may, as is often convenient, suppose that $P_n(x)$ is defined by the equation

$$P_n(x) = \frac{n!\, n^x}{x(x+1)(x+2)\ldots(x+n)}. \quad \ldots\ldots\ldots\ldots(1a)$$

When a distinction is needed, the form (1) may be called the *first* and the form (1a) the *second* form of $P_n(x)$.

First Proof. Take the form (1) and write

$$n!\,n^{x-1}=(n-1)!\,n^x=(n-1)!\left(\frac{2}{1}\cdot\frac{3}{2}\cdots\frac{r+1}{r}\cdots\frac{n}{n-1}\right)^x$$

$$=(n-1)!\prod_{r=1}^{n-1}\left(1+\frac{1}{r}\right)^x;$$

also. $$\frac{(x+1)(x+2)\dots(x+r)\dots(x+n-1)}{(n-1)!}=\prod_{r=1}^{n-1}\left(1+\frac{x}{r}\right).$$

Thus $$xP_n(x)=\prod_{r=1}^{n-1}\left\{\left(1+\frac{1}{r}\right)^x\left(1+\frac{x}{r}\right)^{-1}\right\}=\prod_{1}^{n-1}(f_r).$$

Now $$f_r=\left(1+\frac{x}{r}+\frac{x(x-1)}{2r^2}+\frac{A}{r^3}\right)\left(1-\frac{x}{r}+\frac{x^2}{r^2}-\frac{B}{r^3}\right)$$

$$=1+\tfrac{1}{2}\frac{x(x-1)}{r^2}+\frac{C}{r^3},$$

where A, B, C are all finite; therefore

$$r^2\,|f_r-1|\to\tfrac{1}{2}\,|x(x-1)| \text{ when } r\to\infty$$

so that $\Sigma(f_r-1)$ converges absolutely for every x. Hence $P_n(x)$ converges absolutely when $n\to\infty$.

Second Proof. Taking the definition (1*a*) we may write

$$\frac{1}{xP_n(x)}=\frac{(x+1)(x+2)\dots(x+r)\dots(x+n)}{1\,.\,2\dots r\dots n}\,e^{-x\log n}$$

since $n^{-x}=e^{-x\log n}$. Now express the factor $(x+r)/r$ in the form

$$\left[\left(1+\frac{x}{r}\right)e^{-\frac{x}{r}}\right]e^{\frac{x}{r}}$$

and we find $$\frac{1}{xP_n(x)}=e^{xC_n}\prod_{r=1}^{n}\left[\left(1+\frac{x}{r}\right)e^{-\frac{x}{r}}\right]$$

where $$C_n=1+\tfrac{1}{2}+\tfrac{1}{3}+\dots+\frac{1}{n}-\log n.$$

Now (Exercises II, 8) $C_n\to\gamma$, Euler's Constant, when $n\to\infty$. Also by § 88, Ex. 5, the product $\Pi[(1+x/n)e^{-x/n}]$ converges absolutely for every x; it also converges uniformly since $n^2\,|f_n-1|\to\tfrac{1}{2}\,|x^2|$ and therefore the M-Test applies, for we may take $M_n=\tfrac{1}{2}K^2/n^2$ where K is any given number. Hence $1/xP_n(x)$ converges absolutely and uniformly for every x when $n\to\infty$, so that

$$\frac{1}{\Gamma(x)}=e^{\gamma x}\,x\prod_{n=1}^{\infty}\left[\left(1+\frac{x}{n}\right)e^{-\frac{x}{n}}\right]. \quad\dots\dots\dots\dots\dots(2)$$

Equation (2) may, since $x\Gamma(x)=\Gamma(x+1)$, (see next article) be expressed in the form

$$\Gamma(x+1)=e^{-\gamma x}\prod_{n=1}^{\infty}\left\{\left(1+\frac{x}{n}\right)^{-1}\cdot e^{\frac{x}{n}}\right\}. \quad \ldots\ldots\ldots\ldots (3)$$

Obviously $1/\Gamma(x)$ is zero and $\Gamma(x)$ infinite if x is zero or a negative integer.

The product (2) is usually called *Weierstrass's Form.* The name " Gamma Function " and the notation $\Gamma(x)$ are due to Legendre.

Gauss's Π-*Function.* The function $\Gamma(x+1)$ is the same as the function $\Pi(x)$ introduced by Gauss. As will be seen immediately $\Gamma(x+1)=x\Gamma(x)$ so that $\Pi(x)$ is the limit of $xP_n(x)$ when $n\to\infty$.

Cor. If m is a *fixed* positive integer $\Gamma(x)$ is the limit of $P_{mn}(x)$ for $mn\to\infty$. For, when a sequence $P_n, P_{n+1}, P_{n+2}, \ldots$ tends to a limit any partial sequence selected from it, P_{n+p}, P_{n+q}, P_{n+} $(p<q<r\ldots)$ will tend to the same limit. $P_{mn}, P_{mn+1}, P_{mn+2}, \ldots$ is such a partial sequence.

96. Properties of $\Gamma(x)$. The following properties are easily deduced from the first form of the definition.

(1) $\Gamma(1)=1$.

$P_n(1)=1$ and therefore $\Gamma(1)=1$.

(2) $\Gamma(x+1)=x\Gamma(x)$ or $\Gamma(x)=(x-1)\Gamma(x-1)$.

$$P_n(x+1)=xP_n(x)\cdot\frac{n}{x+n} \text{ so that } \Gamma(x+1)=x\Gamma(x).$$

(3) If $x=n$, a positive integer, $\Gamma(n)=(n-1)!$.

Apply (2) repeatedly.

$$\Gamma(n)=(n-1)\Gamma(n-1)=(n-1)(n-2)\Gamma(n-2)=\ldots$$

the last factor being $\Gamma(1)$ which is unity.

Cor. 1. $\Pi(n)=\Gamma(n+1)=n!$.

Cor. 2. If p is a positive proper fraction and n a positive integer

$$\Gamma(n+p)=(n-1+p)(n-2+p)\ldots(1+p)\Gamma(1+p) \quad \ldots\text{(i)}$$

$$\begin{aligned}\Gamma(1-p)&=-p\Gamma(-p)=-p(-p-1)\Gamma(-p-1)\\&=(-p)(-p-1)\ldots(-p-n)\Gamma(-p-n),\end{aligned}$$

so that $\Gamma(-n-p)=(-1)^{n+1}\Gamma(1-p)/p(p+1)(p+2)\ldots(p+n)$. (ii)

(4) $\Gamma(x)\Gamma(1-x)=\pi/\sin\pi x$.

If x is *real or complex*, but not zero nor a positive or negative integer, the product $P_n(x)P_n(1-x)$ may be expressed as

$$\frac{n!\,n^{x-1}}{x(1+x)(2+x)\dots(n-1+x)}\times\frac{n!\,n^{1-x}\,.\,n^{-1}}{(1-x)(2-x)\dots(n-1-x)(n-x)},$$

and this is equal to unity divided by

$$x(1+x)\left(1+\frac{x}{2}\right)\dots\left(1+\frac{x}{n-1}\right)$$
$$\times(1-x)\left(1-\frac{x}{2}\right)\dots\left(1-\frac{x}{n-1}\right)\left(1-\frac{x}{n}\right)$$

that is, $$\left[x\prod_{r=1}^{n-1}\left(1-\frac{x^2}{r^2}\right)\right]\times\left(1-\frac{x}{n}\right).$$

When $n\to\infty$ the product last written tends to $(\sin\pi x)/\pi$ so that $\Gamma(x)\Gamma(1-x)$ is the reciprocal of $(\sin\pi x)/\pi$. Thus, whether x is real or complex (the values 0 and positive and negative integers excluded) $\Gamma(x)\Gamma(1-x)=\pi/\sin\pi x$.

If $x+y=1$, $\Gamma(x)\Gamma(y)=\pi/\sin\pi x=\pi/\sin\pi y$.

Cor. 1. $\Gamma(\tfrac{1}{2})\Gamma(\tfrac{1}{2})=\pi/\sin\dfrac{\pi}{2}$; $\Gamma(\tfrac{1}{2})=\sqrt{\pi}$, since $\Gamma(\tfrac{1}{2})$ is positive.

Cor. 2. $\Pi(x)\Pi(-x)=\Gamma(1+x)\Gamma(1-x)=\pi x/\sin\pi x$.

Cor. 3. $\Pi(-\tfrac{1}{2})=\Gamma(1-\tfrac{1}{2})=\Gamma(\tfrac{1}{2})=\sqrt{\pi}$.

(5) $\underset{x\to-n}{\mathcal{L}}(x+n)\Gamma(x)=(-1)^n/n!$ (n a positive integer).

For, $(x+n)\Gamma(x)=\Gamma(x+n+1)/x(x+1)(x+2)\dots(x+n-1)$.

(6) If a is real and positive and n a positive integer,

$$\frac{\Gamma(n+a)}{n^a\Gamma(n)}\to 1 \text{ when } n\to\infty.$$

For $$\frac{\Gamma(n+a)}{n^a\Gamma(n)}=\frac{\Gamma(a)\,.\,a(a+1)(a+2)\dots(a+n-1)}{n!\,n^{a-1}}=\frac{\Gamma(a)}{P_n(a)}$$

and $P_n(a)\to\Gamma(a)$ when $n\to\infty$.

97. Gauss's Function $\psi(x)$. If $\psi(x)$ denote the derivative of $\log\Gamma(1+x)$ with respect to x, the function $\psi(x)$ is called *Gauss's Function* $\psi(x)$. A Table of values of $\psi(x)$ is given in No. I of *Tracts for Computers* (Cambridge University Press); $\psi(x)$ is there named the *Digamma Function* and a Table is also given of the values of the function $d\psi(x)/dx$ which is there called the *Trigamma Function.*

It should be noted, however, that the symbol $\psi(x)$ is frequently (in English text-books, usually) taken to be the derivative, not of $\log\Gamma(1+x)$, but of $\log\Gamma(x)$. If, for the sake of distinction, $\psi_1(x)$ is taken to mean $d\log\Gamma(x)/dx$, the relation between $\psi_1(x)$ and Gauss's Function $\psi(x)$ is simply $\psi_1(x)=\psi(x-1)$.

If a is constant $\psi(ax)$ means $d\log\Gamma(1+ax)/d(ax)$; that is

$$\psi(ax)=\frac{1}{a}\frac{d\log\Gamma(1+ax)}{dx}; \quad \psi(-x)=-\frac{d\log\Gamma(1-x)}{dx}.$$

The following properties of Gauss's Function $\psi(x)$ are easily proved by using Weierstrass's form of $\Gamma(1+x)$.

(1) $\psi(x)=-\gamma+\sum_1^\infty\left(\frac{1}{n}-\frac{1}{x+n}\right)$ (2) $\psi(0)=-\gamma$.

(3) $\psi(x+1)=\psi(x)+\frac{1}{x+1}$; $\psi(x+n)=\psi(x)+\sum_{r=1}^{n}\frac{1}{x+r}$.

(4) $\psi(n)=-\gamma+\sum_{r=1}^{n}\frac{1}{r}$.

(5) $\psi(-x-1)-\psi(x)=\pi\cot\pi x$.

(5a) $\psi_1(1-x)-\psi_1(x)=\pi\cot\pi x$ if $\psi_1(x)=\frac{d\log\Gamma(x)}{dx}$.

(6) $\frac{d\,\psi(x)}{dx}=\frac{d^2\log\Gamma(1+x)}{dx^2}=\sum_{n=1}^{\infty}\frac{1}{(x+n)^2}$.

To prove (5) observe that

$$\Gamma(-x)\Gamma(1+x)=\pi/\sin\pi(x+1)=-\pi/\sin\pi x$$

so that $$-\frac{d\log\Gamma(-x)}{dx}-\frac{d\log\Gamma(1+x)}{dx}=\frac{d\log(\sin\pi x)}{dx}=\pi\cot\pi x.$$

Ex. 1. $\psi'(0)=\frac{\pi^2}{6}$.

By (6), $\psi'(0)=\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$, (§ 93, Ex. 3).

Ex. 2. $\psi'(-\frac{1}{2})=\frac{\pi^2}{2}$.

By (6), $\psi'(-\frac{1}{2})=\sum_{n=1}^{\infty}\frac{1}{(n-\frac{1}{2})^2}=4\sum_1^\infty\frac{1}{(2n-1)^2}=\frac{\pi^2}{2}$, (§ 93, Ex. 4).

For other relations see Exercises XII, 8, 9, 10, 14, 15.

98. Examples. The following examples indicate methods of expressing some infinite products as Gamma Functions.

1. $\prod_{n=1}^{\infty} \dfrac{(n+a)(n+b)}{n(n+a+b)} = \dfrac{\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(b+1)}$.

In § 95, (1) let $x=a+1$ and we find

$$(a+1)(a+2)\dots(a+n)=n!\,n^a/P_n(a+1),$$

and therefore, the other factors being treated similarly,

$$\prod_{r=1}^{n} \frac{(r+a)(r+b)}{r(r+a+b)} = \frac{n!\,n^a}{P_n(a+1)} \cdot \frac{n!\,n^b}{P_n(b+1)} \cdot \frac{1}{n!} \cdot \frac{P_n(a+b+1)}{n!\,n^{a+b}}$$
$$= \frac{P_n(a+b+1)}{P_n(a+1)P_n(b+1)}.$$

The result follows at once.

2. $(1-x)(1+\frac{1}{2}x)(1-\frac{1}{3}x)(1+\frac{1}{4}x)\dots = \dfrac{\Gamma(\frac{1}{2})}{\Gamma(1+\frac{1}{2}x)\Gamma(\frac{1}{2}-\frac{1}{2}x)}$.

Let $f_{2n}=(1-x)(1+\frac{1}{2}x)\dots\left(1-\dfrac{x}{2n-1}\right)\left(1+\dfrac{x}{2n}\right)$; then

$f_{2n+1}=f_{2n}[1-x/(2n+1)]$ so that, if f_{2n} tends to a limit, f_{2n+1} tends to the same limit and therefore f_n tends to a limit whether n is even or odd. Now

$$\left(1-\frac{x}{2n-1}\right)\left(1+\frac{x}{2n}\right)=\frac{\left(n-\dfrac{1+x}{2}\right)\left(n+\dfrac{x}{2}\right)}{n(n-\frac{1}{2})},$$

and the result follows from Ex. 1 by taking $a=-\frac{1}{2}(1+x)$, $b=\frac{1}{2}x$.

3. If $\quad u_n=\dfrac{\alpha(\alpha+1)\dots(\alpha+n-1)\,.\,\beta(\beta+1)\dots(\beta+n-1)}{1\,.\,2\,.\dots.\,n\,.\,\gamma(\gamma+1)\dots(\gamma+n-1)}$,

investigate the convergence of Σu_n.

The series Σu_n is $F(\alpha, \beta, \gamma, 1)$, § 60, Ex. 3. Proceeding as in Ex. 1, we see that

$$u_n=\frac{n!\,n^{\alpha-1}}{P_n(\alpha)} \cdot \frac{n!\,n^{\beta-1}}{P_n(\beta)} \cdot \frac{1}{n!} \cdot \frac{P_n(\gamma)}{n!\,n^{\gamma-1}} = \frac{P_n(\gamma)}{P_n(\alpha)P_n(\beta)} \cdot \frac{1}{n^{\gamma-\alpha-\beta+1}}.$$

Suppose that $\Gamma(\alpha)$, $\Gamma(\beta)$ and $\Gamma(\gamma)$ are definite numbers and let α_1, β_1 and γ_1 be the real parts of α, β and γ; then

$$|u_n|=\left|\frac{P_n(\gamma)}{P_n(\alpha)P_n(\beta)}\right|\frac{1}{n^{\gamma_1-\alpha_1-\beta_1+1}} < \frac{K}{n^{\gamma_1-\alpha_1-\beta_1+1}}$$

where K is a constant which, for large values of n, differs little from $|\Gamma(\gamma)/\Gamma(\alpha)\Gamma(\beta)|$. Hence Σu_n converges absolutely if $\gamma_1-\alpha_1-\beta_1$ is positive, that is, $\gamma_1>\alpha_1+\beta_1$.

Again, $$\mathop{\mathcal{L}}_{n\to\infty} |nu_n| = \mathcal{L}\frac{K}{n^{\gamma_1-\alpha_1-\beta_1}}=0,$$

when Σu_n converges absolutely.

4. $\prod_{n=1}^{\infty}\left\{\left(1+\frac{x}{c+n}\right)e^{-\frac{x}{n}}\right\}=\frac{e^{-\gamma x}\Gamma(1+c)}{\Gamma(1+c+x)}$.

We may write

$$\left(1+\frac{x}{c+n}\right)e^{-\frac{x}{n}}=\frac{\left(1+\frac{c+x}{n}\right)e^{-\frac{c+x}{n}}}{\left(1+\frac{c}{n}\right)e^{-\frac{c}{n}}}.$$

But $$\frac{1}{\Gamma(c+x)}=e^{\gamma(c+x)}(c+x)\prod_{n=1}^{\infty}\left\{\left(1+\frac{c+x}{n}\right)e^{-\frac{c+x}{n}}\right\}$$

and $$\frac{1}{\Gamma(c)}=e^{\gamma c}c\prod_{n=1}^{\infty}\left\{\left(1+\frac{c}{n}\right)e^{-\frac{c}{n}}\right\};$$

since $\Gamma(1+c+x)=(c+x)\Gamma(c+x)$, the result follows at once.

The following example gives an interesting illustration of Cesàro's Theorem, § 65.

5. Let $f(x)=1^{\mu-1}x+2^{\mu-1}x^2+3^{\mu-1}x^3+\ldots+n^{\mu-1}x^n+\ldots$; show that, if μ is positive, $(1-x)^\mu f(x)\to\Gamma(\mu)$ when $x\to 1$.

The series converges if $|x|<1$ but diverges if $x=1$. The function $(1-x)^{-\mu}$ is, by the Binomial Theorem, represented by the series

$$\sum_{n=0}^{\infty}\frac{\mu(\mu+1)(\mu+2)\ldots(\mu+n-1)}{n!}x^n,$$

which (by Raabe's Test) diverges if $x=1$. In Cesàro's Theorem take $g(x)=(1-x)^{-\mu}$; then

$$\mathop{\mathcal{L}}_{x\to 1}(1-x)^\mu f(x)=\mathop{\mathcal{L}}_{x\to 1}\frac{f(x)}{g(x)}\mathop{\mathcal{L}}_{n\to\infty}\frac{n!\,n^{\mu-1}}{\mu(\mu+1)\ldots(\mu+n-1)},$$

provided the last limit exists, as it does, being $\Gamma(\mu)$.

99. The Hypergeometric Function. The hypergeometric series, when the real part of $(\gamma-\alpha-\beta)$ is positive, is equal to $F(\alpha,\beta,\gamma,1)$; when the parameters α,β,γ satisfy the condition just stated, the function $F(\alpha,\beta,\gamma,1)$ can be expressed in terms of Gamma Functions, namely

$$F(\alpha,\beta,\gamma,1)=\frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}.$$

The theorem may be verified as follows. Let u_n, v_n and w_n be the coefficients of x^n in the series for $F(\alpha,\beta,\gamma,x)$, $F(\alpha,\beta,\gamma+1,x)$ and $F(\alpha-1,\beta,\gamma,x)$ respectively;

$$u_0=1=v_0=w_0.$$

It is not hard to see that the following relations hold:

(i) $u_n - u_{n+1} = \left(1 - \dfrac{\beta}{\gamma}\right) v_n - w_{n+1}$, $n = 0, 1, 2, \ldots$.

(ii) $(\gamma - \alpha)(u_n - w_n) = \beta u_{n-1} + [(n-1)u_{n-1} - nu_n]$, $n = 1, 2, \ldots$.

All three series Σu_n, Σv_n, Σw_n converge when the real part of $\gamma - \alpha - \beta$ is positive; also (by § 98, Ex. 3) $nu_n \to 0$ when $n \to \infty$.

Now sum (i) from $n = 0$ to $n = \infty$; therefore

$$1 = \left(1 - \frac{\beta}{\gamma}\right) F(\alpha, \beta, \gamma + 1, 1) - [F(\alpha - 1, \beta, \gamma, 1) - 1],$$

or
$$\gamma F(\alpha - 1, \beta, \gamma, 1) = (\gamma - \beta) F(\alpha, \beta, \gamma + 1, 1) \quad \ldots\ldots(a)$$

Again, sum (ii) from $n = 1$ to $n = \infty$; therefore

$$(\gamma - \alpha)\{F(\alpha, \beta, \gamma, 1) - F(\alpha - 1, \beta, \gamma, 1)\} = \beta F(\alpha, \beta, \gamma, 1) - \underset{n \to \infty}{\mathcal{L}}(nu_n)$$
$$= \beta F(\alpha, \beta, \gamma, 1),$$

or
$$(\gamma - \alpha - \beta) F(\alpha, \beta, \gamma, 1) = (\gamma - \alpha) F(\alpha - 1, \beta, \gamma, 1). \quad \ldots(b)$$

Eliminate $F(\alpha - 1, \beta, \gamma, 1)$ between (a) and (b); we thus find

$$F(\alpha, \beta, \gamma, 1) = \frac{(\gamma - \alpha)(\gamma - \beta)}{\gamma(\gamma - \alpha - \beta)} \cdot F(\alpha, \beta, \gamma + 1, 1). \quad \ldots..(c)$$

Now apply the formula (c) repeatedly so as to increase the third element to $\gamma + n$; thus $F(\alpha, \gamma, \beta, 1)$ becomes equal to

$$\frac{(\gamma-\alpha)(\gamma-\alpha+1)\ldots(\gamma-\alpha+n-1) \,.\, (\gamma-\beta)(\gamma-\beta+1)\ldots(\gamma-\beta+n-1)}{\gamma(\gamma+1)\ldots(\gamma+n-1) \,.\, (\gamma-\alpha-\beta)(\gamma-\alpha-\beta+1)\ldots(\gamma-\alpha-\beta+n-1)}$$
$$\times F(\alpha, \beta, \gamma+n, 1).$$

But (see § 98, Examples 1, 3) the coefficient of $F(\alpha, \beta, \gamma + n, 1)$ has as its limit when $n \to \infty$

$$\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)/\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)$$

Again, when $n \to \infty$, every term in the series for

$$F(\alpha, \beta, \gamma + n, 1)$$

tends to zero except the first term which is unity. Hence

$$F(\alpha, \beta, \gamma, 1) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)}.$$

100. Gauss's Formula for $\Gamma(mx)$. If m is a positive integer

$$\Gamma(x)\Gamma\left(x + \frac{1}{m}\right)\Gamma\left(x + \frac{2}{m}\right)\ldots\Gamma\left(x + \frac{m-1}{m}\right) = \frac{(2\pi)^{\frac{m-1}{2}}\Gamma(mx)}{m^{mx-\frac{1}{2}}}$$

Take the first form of $P_n(x)$. The function $P_n\left(x+\frac{r}{m}\right)$, when numerator and denominator have been multiplied by m^n is

$$P_n\left(x+\frac{r}{m}\right)=\frac{m^n \, . \, n! \, . \, n^{x+\frac{r}{m}-1}}{(mx+r)(mx+m+r)\ldots[mx+(n-1)m+r]}.$$

If each of the functions $P_n(x+r/m)$, for $r=0, 1, \ldots m-1$, is expressed in the same way and the m functions then multiplied together, it is readily seen that the denominator will be the product of the factors $(mx+r)$ from $r=0$ to $r=mn-1$ inclusive. The *first* factor from each denominator gives

$$mx(mx+1)\ldots(mx+m-1);$$

then the *second* factor from each gives

$$(mx+m)(mx+m+1)\ldots(mx+2m-1),$$

and so on. Hence

$$\prod_{r=0}^{m-1} P_n\left(x+\frac{r}{m}\right)=\frac{(m^n n!)^m \, . \, n^{mx-\frac{1}{2}(m+1)}}{mx(mx+1)(mx+2)\ldots(mx+mn-1)}. \quad \ldots(1)$$

Now, § 95, Cor., we may take $\Gamma(mx)$ as the limit for mn tending to infinity of $P_{mn}(mx)$, where

$$P_{mn}(mx)=\frac{(mn)! \, (mn)^{mx-1}}{mx(mx+1)(mx+2)\ldots(mx+mn-1)}. \quad (2)$$

If we now divide corresponding members of (1) and (2) we find

$$\frac{\prod_{r=0}^{m-1} P_n\left(x+\frac{r}{m}\right)}{P_{mn}(mx)} \, . \, m^{mx-1}=\frac{(m^n n!)^m}{(mn)! \, n^{\frac{m-1}{2}}}. \quad (3)$$

and the expression on the right is independent of x; its limit for $n\to\infty$ is easily found by using the value of $n!$ in Exercises II, 30 to be $(2\pi)^{\frac{m-1}{2}} \, . \, m^{-\frac{1}{2}}$. The limit for $n\to\infty$ of the left hand side is $\Gamma(x)\Gamma(x+1/m)\ldots\Gamma[x+(m-1)/m] \, . \, m^{mx-1}/\Gamma(mx)$ so that the formula is established.

The student might work out independently, by the same method, the particular case $\Gamma(x)\Gamma(x+\frac{1}{2})$, obtained by a totally different method in the *Elementary Treatise*, p. 450.

In Gauss's Formula put $x+1/m$ in place of x; then

$$\Gamma(mx+1)=(2\pi)^{\frac{1-m}{2}} \, . \, m^{mx+\frac{1}{2}}\prod_{s=0}^{m-1}\Gamma\left(1+x-\frac{s}{m}\right). \quad \ldots\ldots..(4)$$

Now $$\psi(mx)=\frac{d\log\Gamma(mx+1)}{d(mx)}=\frac{1}{m}\frac{d\log\Gamma(mx+1)}{dx}$$

and therefore, by taking logarithms and differentiating (4) we find

$$\psi(mx)=\log m+\frac{1}{m}\sum_{s=0}^{m-1}\psi\left(x-\frac{s}{m}\right). \quad \ldots\ldots\ldots\ldots(5)$$

EXERCISES XII.

1. If $f(x)=\prod_{n=1}^{\infty}\left\{\left(1+\frac{x}{n\pi}\right)e^{-\frac{x}{n\pi}}\right\}$, prove that $f(x)\times f(-x)=\sin x/x$.

2.* If $\varphi(x)=\prod_{r=-m}^{n}{}'\left(1-\frac{x}{r\pi}\right)$, show that when m and n tend to infinity $\varphi(x)$ tends to the limit $(\sin x/x)a^{x/\pi}$ where a is the limit of (m/n).

Note that $$\prod_{r=1}^{n}\left(1-\frac{x}{r\pi}\right)=\prod_{r=1}^{n}\left\{\left(1-\frac{x}{r\pi}\right)e^{\frac{x}{r\pi}}\right\}\times e^{s_n},$$

where $$s_n=-\frac{x}{\pi}\left\{\left(1+\tfrac{1}{2}+\ldots+\frac{1}{n}\right)-\log n\right\}-\frac{x}{\pi}\log n.$$

Thus $\varphi(x)$ tends to $\sin x/x$ if, and only if, $a=1$, that is, if m and n tend to infinity "in a ratio of equality."

3. (i) $$\sin\pi x=\pi x\prod_{n=-\infty}^{\infty}{}'\left\{\left(1-\frac{x}{n}\right)e^{\frac{x}{n}}\right\};$$

(ii) $$\pi\cot\pi x=\frac{1}{x}+\sum_{n=-\infty}^{\infty}{}'\left\{\frac{1}{n}+\frac{1}{x-n}\right\};$$

(iii) $$\frac{\pi^2}{n^2\,\pi x}=\sum_{n=-\infty}^{\infty}\frac{1}{(x-n)^2}.$$

4. Prove the following statements:

(i) $$f(x)=\prod_{n=-\infty}^{\infty}\left\{\left(1-\frac{x}{n-a}\right)e^{\frac{x}{n-a}}\right\}=\frac{\sin\pi(x+a)}{\sin\pi a}e^{-\pi x\cot\pi a};$$

(ii) $$\sum_{n=-\infty}^{\infty}\left\{\frac{1}{n-a}+\frac{1}{x+a-n}\right\}=\pi\{\cot\pi(x+a)-\cot\pi a\}.$$

* The symbol Π', with the accent on Π, indicates that the value $r=0$ is excluded so that r takes the values $-m$, $-(m-1)$, ... -2, -1, 1, 2, ... n. A similar meaning is assigned to the symbol of summation Σ'.

5. If $\varphi(x)=x\prod_{n=-\infty}^{\infty}{}'\left\{\left(1-\frac{x}{n}\right)e^{\frac{x}{n}}\right\}$ and if $f(x)$ is the function defined by the product in Ex. 4, (i), prove, without using the circular functions and simply by transforming the products, that

(i) $f(x)=\frac{\varphi(x+a)}{\varphi(a)}e^{-Ax}$, $A=\frac{\varphi'(a)}{\varphi(a)}$;

(ii) $\varphi(x+1)=-\varphi(x)$, $\varphi'(x+1)/\varphi(x+1)=\varphi'(x)/\varphi(x)$;

(iii) $f(x+1)=-e^{-A}f(x)$.

6. $\prod_{n=0}^{\infty}\left\{\left(1+\frac{c}{n+a}\right)\left(1-\frac{c}{n+b}\right)\right\}=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+c)\Gamma(b-c)}$.

Show that if $a=\gamma$, $b=\gamma-\alpha-\beta$, $c=-\alpha$, where the real part of $(\gamma-\alpha-\beta)$ is positive, the infinite product is $F(\alpha,\beta,\gamma,1)$.

7. $\underset{n\to\infty}{\mathcal{L}}\prod_{r=1}^{n-1}\left(1-\frac{r}{n}\right)^{\frac{1}{r}}=e^{-\frac{\pi^2}{6}}$.

8. $\prod_{n=0}^{\infty}\left\{\left(1+\frac{a}{x+n}\right)e^{-\frac{a}{x+n}}\right\}=\frac{\Gamma(x)}{\Gamma(x+a)}e^{a\psi(x-1)}$. (Mellin.)

9. (i) $\psi(\frac{1}{2})+\gamma=2-2\log 2$; (ii) $\psi(-\frac{1}{4})-\psi(-\frac{3}{4})=\pi$.

10. (i) $\sum_{n=0}^{\infty}\left\{\frac{1}{n+1}-\log\left(1+\frac{1}{x+n}\right)\right\}=\gamma+\log x$;

(ii) $\sum_{n=0}^{\infty}\left\{\log\left(1+\frac{1}{x+n}\right)-\frac{1}{x+n+1}\right\}=\psi(x)-\log x$.

11. If n, p, x are all positive integers, prove that

$$\underset{x\to\infty}{\mathcal{L}}\frac{(x+n)(x+n+1)\dots(x+n+p-1)}{(x+1)(x+2)\dots(x+p)}=1.$$

12. $\underset{n\to\infty}{\mathcal{L}}\frac{x(x+1)(x+2)\dots(x+2n-1)}{1.3.5\dots.(2n-1).2x(2x+2)\dots(2x+2n-2)}=2^{x-1}$.

13. Prove that if $|x|<1$,

(i) $\log\Gamma(1+x)=-\gamma x+\sum_{n=2}(-1)^n S_n\frac{x^n}{n}$

where $$S_n=\frac{1}{1^n}+\frac{1}{2^n}+\frac{1}{3^n}+\dots;$$

(ii) $\log\Gamma(1+x)+\log(1+x)=(1-\gamma)x+\sum_{n=2}(-1)^{n-1}(1-S_n)\frac{x^n}{n}$

(iii) $\log\Gamma(1-x)+\log(1-x)=-(1-\gamma)x-\sum_{n=2}^{\infty}(1-S_n)\frac{x^n}{n}$;

(iv) $\log\Gamma(1+x)-\log\Gamma(1-x)=\log\dfrac{1-x}{1+x}$

$$+2(1-\gamma)x+2\sum_{n=1}^{\infty}(1-S_{2n+1})\frac{x^{2n+1}}{2n+1};$$

(v) $\log\Gamma(1+x)+\log\Gamma(1-x)=\log(\pi x/\sin\pi x)$;

(vi) $\log\Gamma(1+x)=\frac{1}{2}\log\left(\dfrac{\pi x}{\sin\pi x}\right)+\frac{1}{2}\log\dfrac{1-x}{1+x}+(1-\gamma)x$

$$\sum_{n=1}^{\infty}(1-S_{2n+1})\frac{x^{2n+1}}{2n+1}.$$

[To obtain (i) use Weierstrass's form for $\Gamma(1+x)$, expand each logarithm and show that the series may be deranged and expressed in powers of x. The equations (ii) ... (vi) then follow easily.]

14. Deduce by putting $x=-\frac{1}{2}$ in Ex. 13, (vi) that

$$\gamma=1-\log\tfrac{3}{2}-\sum\frac{S_{2n+1}-1}{2n+1}\cdot\frac{1}{2^{2n}}.$$

15. Deduce series for $\psi(x)$ from Ex. 13; for example,

$$\psi(x)=-\gamma+\sum_{n=2}^{\infty}(-1)^n S_n x^{n-1},\quad |x|<1.$$

CHAPTER IX

INTEGRATION OF BOUNDED FUNCTIONS

101. Intervals. Sets. In establishing general theorems in integration there is frequent use of the division of an interval into sub-intervals and of sums of terms associated with sub-intervals; it will obviate inconvenient interruptions in exposition and perhaps emphasize the essential elements in the discussion if we begin with some definitions and explanations.

Numbers $x_1, x_2, \dots, x_{n-1}$ chosen so that

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < b = x_n$$

effect a division, D say, of the interval (a, b) into n sub-intervals, the notation $a = x_0$, $b = x_n$ being adopted for symmetry and for use in summations.

If in one, more or all of the sub-intervals of D there be inserted one or more numbers, a new division, D' say, of the interval (a, b) is made which is said to be *consecutive* to D; the numbers $x_1, x_2, \dots, x_{n-1}$ and the numbers that have been inserted are considered as a single set and are always supposed to be arranged in order of magnitude from a to b. For example D' might be

$$a,\ x_1,\ \xi_1,\ \xi_2,\ x_2,\ x_3,\ x_4,\ \xi_3,\ \dots,\ \xi_p,\ x_{n-1},\ b,$$

where $\xi_1, \xi_2, \dots$ are the numbers inserted and $a < x_1 < \xi_1 < \xi_2 < x_2 \dots < b$.

If there are two different divisions of (a, b), say

$$D_1\,[a, x_1, x_2, \dots, x_{m-1}, b]$$

with m sub-intervals and $D_2\,[a, \xi_1, \xi_2, \dots, \xi_{n-1}, b]$ with n sub-intervals, the division, D_3 say, formed by taking the numbers x and ξ in order of magnitude from a to b is said to be made by *superposition* of the divisions D_1 and D_2. In the division D_3 there will lie between a and b *at most* $(m+n-2)$ numbers; but there may be fewer, since a number x may be equal to a number ξ and every equality of this kind reduces $(m+n-2)$. Thus if $x_2 = \xi_3$, $x_7 = \xi_{12}$ the four numbers $x_2, x_7, \xi_3, \xi_{12}$ would give only two different numbers and the number $(m+n-2)$ would be reduced by 2. It is obvious that D_3 may be considered as *consecutive* both to D_1 and to D_2; if D_3 is taken to be consecutive to D_1

there are not more than $(n-1)$ numbers inserted in the sub-intervals of D_1 but there may be fewer owing to equalities such as $x_2=\xi_3$, $x_7=\xi_{12}$.

The formation of a consecutive division by superposition is of constant occurrence and the student should be quite clear what exactly is being done.

Again, there may be associated with a sub-interval (x_r, x_{r+1}) of a division D a number u_r and with the division D there may be associated a sum v, where

$$v=u_0(x_1-a)+u_1(x_2-x_1)+\ldots+u_r(x_{r+1}-x_r)+\ldots+u_{n-1}(b-x_{n-1})$$
$$=\sum_{r=0}^{n-1}u_r(x_{r+1}-x_r).$$

If the division D varies so will the number v, and if we suppose, as will usually happen, that n may be as large as we please and that the numbers x_r may be chosen arbitrarily so long as they are arranged in order of magnitude from a to b, the numbers v will form an infinite set (v). The properties of the bounds of the set (v) will then be discussed. It is the properties of such sets that lead to the conditions for the existence of an integral.

If the student turn to p. 324 of the *Elementary Treatise* he will see that v is the sum of the expression (1) on that page when $u_r=F(x_r)$; the upper bound of the set (v) is the area $ABDC$ under the curve (Fig. 75). In the absence of theorems on the existence of bounds of infinite sets appeal was made to the conception of an area to determine the limit of the sum (1), p. 324, and establish the existence of the integral. From our present standpoint the process is reversed; the existence of the integral is first established without appeal to geometrical considerations and then the area is defined by an integral.

102. The Sums S and s. Let the function $F(x)$ be single-valued and bounded for the range $a\leq x\leq b$; at present no other restriction, such as continuity, is imposed on $F(x)$.*

Let $D[a, x_1, \ldots, x_{n-1}, b]$ be a division of the interval (a, b) and let $(x_{r+1}-x_r)=h_r$, a positive number that measures the length of the sub-interval (x_r, x_{r+1}). Further, denote by M, m and M_r, m_r the upper and lower bounds respectively of $F(x)$ in the whole interval (a, b) and in the sub-interval (x_r, x_{r+1}).

Now consider the sums S and s where

$$S=M_0h_0+M_1h_1+\ldots+M_{n-1}h_{n-1}=\sum_{r=0}^{n-1}M_rh_r \quad\ldots\ldots\ldots\ldots\ldots(1)$$

$$s=m_0h_0+m_1h_1+\ldots+m_{n-1}h_{n-1}=\sum_{r=0}^{n-1}m_rh_r \quad\ldots\ldots\ldots(2)$$

* For the value of $F(x)$ at a point of discontinuity in (a, b) see § 29.

S and s are called respectively the *upper* and *lower* sums for the function $F(x)$ and the division D of the interval (a, b).

In whatever way the numbers x_r are chosen we have

$$m \leqq m_r \leqq M_r \leqq M,$$

and therefore, from equations (1) and (2),

$$m(b-a) \leqq s \leqq S \leqq M(b-a) \quad \dots\dots\dots\dots\dots\dots(3)$$

since $$h_0 + h_1 + \dots + h_{n-1} = b - a.$$

From the inequalities (3) take the following

$$s \leqq S, \quad S \geqq m(b-a), \quad s \leqq M(b-a) \quad \dots\dots\dots\dots(3a)$$

Thus it is seen that *for the same division* of (a, b), $s \leqq S$. Again, the inequalities $S \geqq m(b-a)$ and $s \leqq M(b-a)$ are true in whatever way the division D may be varied. The variation of D may be made in an infinite number of ways by varying n and the numbers x_r and for each division there is a corresponding S and a corresponding s. Hence the inequalities (3a) for S and s show that the set (S) has a lower bound, L say, and the set (s) an upper bound, which may be called l.

Some properties of S and s will now be proved.

1. The inequalities

$$m(b-a) \leqq S \leqq M(b-a), \quad m(b-a) \leqq s \leqq M(b-a)$$

are true whatever be the division of (a, b)

These inequalities are obvious but important.

2. If the division D_1 of (a, b) is *consecutive* to the division D, and if S_1 and s_1 are the upper and lower sums respectively for the division D_1, then $S \geqq S_1$, $s \leqq s_1$; that is, in passing from any division to a *consecutive* division the sum S decreases or is stationary while the sum s increases or is stationary.

Suppose first that D_1 contains only one number ξ, where $x_r < \xi < x_{r+1}$, that does not occur in D, and let M', M'' be the upper bounds of $F(x)$ in the intervals (x_r, ξ), (ξ, x_{r+1}) respectively. All the terms in S and S_1 are the same except that instead of the term $M_r(x_{r+1} - x_r)$ in S there is in S_1 the sum of the two terms $M'(\xi - x_r)$ and $M''(x_{r+1} - \xi)$. Hence

$$S - S_1 = M_r(x_{r+1} - x_r) - [M'(\xi - x_r) + M''(x_{r+1} - \xi)];$$

but $(x_{r+1} - x_r) = (\xi - x_r) + (x_{r+1} - \xi)$ and therefore

$$S - S_1 = (M_r - M')(\xi - x_r) + (M_r - M'')(x_{r+1} - \xi). \quad \dots\dots(4)$$

One of the numbers M', M'' is equal to M_r while the other may be less than M_r or equal to M_r; each term on the right of (4) is therefore zero or positive so that $S - S_1$ is zero or positive, that is, $S \geqq S_1$.

In a similar way it may be seen that $s \leqq s_1$ since the lower bounds, m', m'' say, of $F(x)$ in the intervals (x_r, ξ), (ξ, x_{r+1}) are greater than or equal to m_r.

Suppose next that D_1 contains more than one number that does not occur in D; these numbers may be supposed to be inserted in succession and since at each insertion S decreases and s increases (if there is any change at all) we see that however many numbers there may be in D_1 that do not occur in D the relations $S \geqq S_1$ and $s \leqq s_1$ are true.

3. If the consecutive division D_1 contains μ numbers that do not occur in D and if for the division D the length (h_r) of each interval is less than h then

$$0 \leqq S - S_1 < \mu(M - m)h, \; 0 \leqq s_1 - s < \mu(M - m)h.$$

The relations $0 \leqq S - S_1$, $0 \leqq s_1 - s$ have been proved in 2.

As before, suppose first that only one point ξ, where $x_r < \xi < x_{r+1}$, has been inserted in D. Then, $M_r - M' \leqq M - m$, $M_r - M'' \leqq M - m$ and therefore by equation (4) we find

$$S - S_1 \leqq (M - m)(x_{r+1} - x_r) < (M - m)h.$$

Thus the decrease in S due to the insertion of one number is less than $(M - m)h$. Suppose next that μ numbers have been inserted. As before, they may be supposed to be inserted in succession, and as the insertion of each additional number produces a change that is less than $(M - m)h$ the insertion of μ numbers produces a change that is less than $\mu(M - m)h$ so that $S - S_1$ is less than $\mu(M - m)h$.

In the same way the relation $s_1 - s < \mu(M - m)h$ is proved; we thus find the inequalities stated.

4. The lower sum s for any one division cannot exceed the upper sum S for any other division and not merely for the same division (as shown in (3*a*)).

Let S, s and S', s' be, respectively, the two sums for two different divisions D and D' of (a, b) and let S'', s'' be the sums for the division D'' formed by the superposition of D and D'.

If D'' be taken to be consecutive to D we have (by 2)

$$S \geqq S'', \quad s \leqq s'',$$

while if taken to be consecutive to D' we have

$$S' \geqq S'', \quad s' \leqq s''.$$

But $s'' \leqq S''$ since both sums belong to the division D''; therefore $\quad s \leqq s'' \leqq S'' \leqq S' \quad$ and $\quad s' \leqq s'' \leqq S'' \leqq S$

as was to be proved.

5. $l \leqq L$. By the definitions of the bounds l and L there are divisions D' and D'' of (a, b) such that s' differs from l and S'' from L by as little as we please; the inequality $l > L$ would therefore imply an s' and an S'' for which $s' > S''$, which has been seen to be impossible.

103. Darboux's Theorem. The following Theorem or Lemma, known as Darboux's Theorem, is of fundamental importance; ε denotes as usual an arbitrarily small positive number and is to be understood in this sense throughout the discussion. The Theorem will be stated in two forms.

First Form. To any given ε there corresponds a positive number h such that $S < L + \varepsilon$ and $s > l - \varepsilon$ for every division of (a, b) in which the length of each (or the longest) sub-interval is less than h.

Second Form. S and s tend to L and l respectively if the number of sub-intervals in the division of (a, b) tends to infinity in such a way that the length of each (or the longest) sub-interval tends to zero.

I. Consider, for example, the sum S. Let D be any division of (a, b) into n sub-intervals such that the length of each is less than h; for the present h is simply a fixed positive number.

Next, since L is the lower bound of the set (S) there is, by the definition of L, a division D' of (a, b)—which will be supposed to contain μ numbers between a and b—such that the corresponding sum S' satisfies the inequality

$$S' < L + \tfrac{1}{2}\varepsilon \quad \text{.............................(1)}$$

Now superpose the divisions D and D' to form a new division D'', and let S'' be the sum for the division D''. If D'' is considered as consecutive to D we have (§ 102, 3)

$$S < S'' + \mu(M - m)h \quad \text{.......................(2)}$$

since there are μ (or fewer) numbers that enter the sub-intervals of D from D'. On the other hand if D'' is taken to be consecutive to D' we have (§ 102, 2 and (1) above)

$$S'' < S' < L + \tfrac{1}{2}\varepsilon \quad \text{.........................(3)}$$

and therefore, by (2),

$$S < L + \tfrac{1}{2}\varepsilon + \mu(M - m)h. \quad \text{...................(4)}$$

The numbers μ, M, m are fixed (though μ depends on ε); we now suppose h chosen so that $\mu(M-m)h < \frac{1}{2}\varepsilon$. Hence, applying (4), we have found h so that $S < L + \varepsilon$; the only restriction on the division D is that the length of each sub-interval is less than h.

The proof for the lower sum s may be left to the student since it follows the same lines as that just given and requires little more than verbal changes.

II. Again $S \geqq L$ since L is the lower bound of the set (S); combining this relation with the inequality $S < L + \varepsilon$, we see that if in any division of (a, b) the number of sub-intervals is so great and at the same time the length of each sub-interval so small that the longest is less than h

$$0 \leqq S - L < \varepsilon.$$

Therefore L is the limit of S under the conditions stated. Obviously we also have $0 \leqq l - s < \varepsilon$ and l is the limit of s.

Ex. 1. Prove that if $F(x)$ is continuous $l = L$.

The proof of Darboux's Theorem requires the property 3 of § 102, but when $F(x)$ is continuous that property need not be appealed to.

Since $F(x)$ is continuous it is *uniformly* continuous and therefore we can choose h so that $(M_r - m_r)$ shall be less than $\varepsilon/(b-a)$ provided that $(x_{r+1} - x_r) = h_r < h$ where r is any of the numbers $0, 1, 2, \ldots, (n-1)$. From equations (1) and (2) of § 102 we find

$$S - s = \sum_{r=0}^{n-1}(M_r - m_r)h_r < \frac{\varepsilon \Sigma h_r}{b-a} = \varepsilon$$

since $\Sigma h_r = (b-a)$.

Next we have the identity

$$S - s = (S - L) + (L - l) + (l - s).$$

Each of the numbers $(S-L)$, $(L-l)$, $(l-s)$ is positive when not zero, and therefore each is less than $S - s < \varepsilon$. But l, L are constants so that $l = L$. Again $S \to L$ and $l \to s$ when each $h_r \to 0$.

Ex. 2. $D[a, x_1, x_2, \ldots, x_{n-1}, b]$ is a division of the interval (a, b); $F(x)$ is bounded in (a, b) and $\varphi(x)$ is continuous and steadily increases

as x increases from a to b. If M_r and m_r are the upper and lower bounds of $F(x)$ in (x_r, x_{r+1}) and if S_ϕ and s_ϕ denote the sums

$$S_\phi = \sum_{r=0}^{n-1} M_r[\varphi(x_{r+1}) - \varphi(x_r)], \quad s_\phi = \sum_{r=0}^{n-1} m_r[\varphi(x_{r+1}) - \varphi(x_r)],$$

show that if n tends to infinity in such a way that the length of each sub-interval (x_r, x_{r+1}) tends to zero S_ϕ and s_ϕ will tend to limits L and l respectively. If $F(x)$ is continuous in (a, b) prove that $l = L$.

(Goursat.)

104. Functions with Limited Variation. A function $F(x)$ which, for a range $a \leqq x \leqq b$, has the property called *limited variation* can be expressed as the difference of two functions which are each positive, monotonic, increasing (or, at least not decreasing) as x increases; for *variation* the word *fluctuation* is sometimes used. This class of functions is important in many investigations and we therefore make a brief reference to them.

Let $D[a, x_1, x_2, \dots, x_{n-1}, b]$ be a division of the interval (a, b) and $F(x)$ a function that is single-valued and bounded in the interval. Consider the differences $\{F(x_{r+1}) - F(x_r)\}$ and also their absolute values $|F(x_{r+1}) - F(x_r)|$; if $v(a, b)$ is the sum of the absolute values we have

$$\left.\begin{aligned} v(a, b) = |F(x_1) - F(a)| + |F(x_2) - F(x_1)| + \dots \\ + |F(b) - F(x_{n-1})| \end{aligned}\right\} \quad \dots.(1)$$

and, identically, for the sum of the differences,

$$\left.\begin{aligned} F(b) - F(a) = \{F(x_1) - F(a)\} + \{F(x_2) - F(x_1)\} + \dots \\ + \{F(b) - F(x_{n-1})\} \end{aligned}\right\} \quad \dots.(2)$$

so that $$v(a, b) \geqq |F(b) - F(a)|.$$

The number $v(a, b)$ is called the *variation* of $F(x)$ in the interval (a, b) for the division D; it is usually different from $|F(b) - F(a)|$. If the set (v) is bounded when the division D varies in all possible ways its upper bound, which will be denoted by $V(a, b)$, is called the *total* variation of $F(x)$ in (a, b) and $F(x)$ is said to be a *function with limited variation in* (a, b)—more fully, with limited *total* variation in (a, b).

If $p(a, b)$ is the sum of the differences in (2) which are positive and $-n(a, b)$ the sum of those which are negative, we have obviously $$F(b) - F(a) = p - n, \quad v = p + n$$
and therefore
$$v = 2p + F(a) - F(b), \quad v = 2n - F(a) + F(b) \quad \dots\dots\dots(3)$$

From (3) it follows that $p(a, b)$ and $n(a, b)$ have upper bounds when v has an upper bound; these are denoted by $P(a, b)$ and $N(a, b)$ respectively and satisfy the equations

$$\left.\begin{aligned} V(a, b) &= 2P(a, b) + F(a) - F(b), \\ V(a, b) &= 2N(a, b) - F(a) + F(b) \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots (4)$$

$P(a, b)$ and $-N(a, b)$ are the *total* positive and negative variations.

The variations $V(a, x)$, $P(a, x)$, $N(a, x)$ in the interval (a, x) where $a \leqq x \leqq b$ are obtained by substituting x for b in the formulae; obviously these numbers are positive and all increase (at least do not decrease) as x increases. In equations (4) put x for b and eliminate $V(a, x)$; then

$$F(x) = F(a) + P(a, x) - N(a, x);$$

take any constant $C > | F(a) |$ and we find

$$F(x) = \{C + F(a) + P(a, x)\} - \{C + N(a, x)\} = \varphi(x) - \psi(x).$$

The functions $\varphi(x)$ and $\psi(x)$ are positive, monotonic, increasing functions; $F(x)$ is thus expressed in the form stated.

It may be remarked that $F(x)$ may be continuous and yet *not* have limited (total) variation (or fluctuation). See Goursat, *Cours d'Analyse*, I, p. 23 (2nd Ed.).

Ex. 1. If $F(x)$ is monotonic for $a \leqq x \leqq b$, $V(a, b) = | F(a) - F(b) |$.

Ex. 2. If $F(x)$ has limited variation in (a, b) and if $a < x < b$ the limits for $h \to 0$ of $F(x+h)$ and $F(x-h)$ are definite numbers.

Ex. 3. If $F(x)$ and $f(x)$ have each limited variation in (a, b) so has their product.

Write the difference $F(x_{r+1}) f(x_{r+1}) - F(x_r) f(x_r)$ in the form

$$F(x_{r+1})\{f(x_{r+1}) - f(x_r)\} + f(x_r)\{F(x_{r+1}) - F(x_r)\};$$

if the upper bounds of $|F(x)|$ and $|f(x)|$ are A, B respectively, the absolute value of the difference is less than or equal to

$$A\,|f(x_{r+1}) - f(x_r)| + B|F(x_{r+1}) - F(x_r)|,$$

and therefore the (total) variation of the product cannot exceed

$$AV(a, b)]_f + BV(a, b)]_F.$$

105. The Definite Integral. The theory now to be explained is usually called *Riemann's Theory of Integration*; other theories, such as that of Lebesgue, will not be discussed as they involve considerations that are outside our limits. We shall suppose the student to be familiar with the terminology and the subject-matter of integration as presented in the *Elementary*

Treatise and shall direct attention chiefly to those aspects of the subject that depend on the more fully developed theory of number and limits given in the preceding pages.

The bounds L and l respectively of the upper and lower sums S and s associated with the *bounded* function $F(x)$ and the interval (a, b) are shown by Darboux's Theorem to be the *limits* of the respective sums, in whatever way the interval (a, b) may be divided provided the number of sub-intervals tends to infinity in such a way that the length of each sub-interval tends to zero.

Definition 1. The limits L and l are defined to be *the upper and lower integrals* respectively of $F(x)$ over the range $a \leqq x \leqq b$ and are expressed by the symbols

$$L = \overline{\int_a^b} F(x)\,dx, \quad l = \underline{\int_a^b} F(x)\,dx.$$

The numbers a and b are called the lower and upper limits respectively of the integrals.

These two integrals always exist, in virtue of Darboux's Theorem, when $F(x)$ is a single-valued, bounded function in (a, b). In general L and l are unequal, but if $L = l$, the common limit is defined to be *the integral* of $F(x)$ for the range $a \leqq x \leqq b$.

Definition 2. If $L = l$ this common limit is defined to be the (definite) integral of $F(x)$ over the range $a \leqq x \leqq b$ and is expressed by the usual notation

$$\int_a^b F(x)\,dx.$$

When $L = l$ the function $F(x)$ is said to be *integrable* over the interval (a, b) and the next step is to state the condition that $F(x)$ should be integrable. It must be remembered that the function $F(x)$ is restricted to being *single-valued and bounded*, and the limits a, b finite; at a later stage (Chapter XIII.) the definition will be extended to cases in which $F(x)$ is not bounded and the limits a and b (one or both) infinite.

106. Condition of Integrability. The condition follows from Darboux's Theorem, and will be stated in the two corresponding forms.

First Form. Given ε (as usual) there must be a positive number h such that $S - s$ will be less than ε for every division

of (a, b) in which the length of each (or the longest) sub-interval is less than h.

Second Form. The difference $S - s$ must tend to zero when the number of sub-intervals in the division of (a, b) tends to infinity in such a way that the length of each (or the longest) sub-interval tends to zero.

Consider the *First Form*. The condition is *necessary*. Let h_p be the longest of the sub-intervals in the division of (a, b); then, by Darboux's Theorem, there is a positive number h such that
$$S - L < \tfrac{1}{2}\varepsilon,\ l - s < \tfrac{1}{2}\varepsilon \text{ if } h_p < h,$$
and therefore if $L = l$,
$$S - s = (S - L) + (l - s) < \varepsilon \text{ if } h_p < h.$$

Again the condition is *sufficient*; because
$$S \geqq L \geqq l \geqq s;\quad L - l \leqq S - s.$$
Hence if $S - s < \varepsilon$ when $h_p < h$ so is $L - l$ and therefore $L = l$, since L and l are constants.

Cor. It may be noted that $F(x)$ is integrable over (a, b) provided there is *one* division of (a, b) for which $S - s < \varepsilon$. The condition is *sufficient* for $L - l \leqq S - s < \varepsilon$ and therefore $L = l$. It is also *necessary* for, as has been seen, if $L = l$ *every* division for which $h_p < h$ makes $S - s < \varepsilon$.

The proof for the *Second Form* is equally simple. If $L = l$ then $S - s$, that is, $(S - L) + (l - s)$ tends to zero when $h_p \to 0$ since in this case $S - L \to 0$ and $l - s \to 0$. Again, if $S - s \to 0$ when $h_p \to 0$ we must have $L = l$ since $L - l \leqq S - s$.

A third form of the condition of integrability may be given which depends on the oscillation ω_r of $F(x)$ in the sub-interval (x_r, x_{r+1}); if M_r and m_r are the upper and lower bounds of $F(x)$ in the sub-interval $\omega_r = M_r - m_r$ (§ 27). In terms of ω_r we have
$$S - s = \Sigma(M_r - m_r)h_r = \Sigma\omega_r h_r.$$

If $F(x)$ is continuous for $a \leqq x \leqq b$ it is possible to choose h (§ 28) so that $\omega_r < \varepsilon/(b - a)$ provided $h_r < h (r = 0, 1, 2, \dots, n - 1)$; in this case $S - s < \varepsilon$ if h_r is less than h, since $\Sigma\omega_r h_r$ is less than $(\Sigma h_r)\varepsilon/(b - a)$ or ε. Thus *a continuous function is always integrable.*

On the other hand if $F(x)$ is discontinuous for $x = c_r$ where $x_r < c_r < x_{r+1}$ the oscillation ω_r of $F(x)$ in the sub-interval (x_r, x_{r+1}) is finite and does not tend to zero when $h_r \to 0$. The

value of $F(x)$ for $x = c_r$ is frequently not determined by the analytical expression for $F(x)$ but it is always assumed to be a fixed number, C_r say, such that $m \leqq C_r \leqq M$, where M and m are the upper and lower bounds of $F(x)$ in (a, b) (or, $m_r \leqq C_r \leqq M_r$. The precise value of C_r is of no importance provided $|C_r|$ is finite.) See § 29. The discontinuity of $F(x)$ at c_r is measured by the lower bound of $(M_r - m_r)$ when x_r and x_{r+1} tend each to c_r and therefore cannot exceed $M - m$.

The condition of integrability may now be stated in another way.

Third Form. The necessary and sufficient condition that the bounded function $F(x)$ be integrable over (a, b) is that to every pair of arbitrarily small positive numbers ω and η there shall correspond a division of (a, b) such that the sum of the lengths of the sub-intervals in which the oscillation of $F(x)$ is greater than or equal to ω will be less than η.

Let δ_r and k_r be the lengths of typical sub-intervals of the division D of (a, b) in which the oscillations ω_r of $F(x)$ are respectively greater than or equal to ω and less than ω; also let $\lambda = \Sigma\delta_r$. Then

$$S - s = \Sigma\omega_r\delta_r + \Sigma\omega_r k_r.$$

Obviously $\quad \Sigma\omega_r\delta_r \geqq \lambda\omega$ but $\Sigma\omega_r\delta_r \leqq \lambda(M - m)$

since $\omega_r \leqq M - m$; further, $\Sigma\omega_r k_r < (b - a)\omega$ since $\Sigma k_r \leqq b - a$.

Hence $$\lambda\omega \leqq S - s \leqq \lambda(M - m) + (b - a)\omega.$$

The condition is necessary; for if ω and η are given $S - s$, which is not less than $\lambda\omega$, cannot be less than $\eta\omega$ unless $\lambda < \eta$ and therefore, if $\eta\omega = \varepsilon$, cannot be less than ε unless $\lambda < \eta$.

The condition is sufficient; for, given ε, the numbers ω and η may be chosen so that

$$\omega = \tfrac{1}{2}\varepsilon/(b - a) \text{ and } \eta = \tfrac{1}{2}\varepsilon/(M - m),$$

and therefore, if $\lambda < \eta$, the division D is such that $S - s < \varepsilon$—an inequality which secures that $L = l$.

107. Other Forms of the Definition of an Integral. Let the notation for the division of the interval (a, b) be the same as in the preceding articles, $F(x)$ being integrable over (a, b). Now take ξ_r so that

$$x_r \leqq \xi_r \leqq x_{r+1},\ r = 0, 1, 2, \ldots, (n - 1), \quad x_{r+1} - x_r = h_r,$$

and consider the sum S_n where

$$S_n = \sum_{r=}^{n-1} F(\xi_r) h_r. \quad \text{.........................(1)}$$

We have $m_r \leqq F(\xi_r) \leqq M_r$ and therefore $s \leqq S_n \leqq S$. Hence

$$I = \mathop{\mathcal{L}}_{n \to \infty} S_n = \int_a^b F(x)\,dx \quad \text{....................}(D_1)$$

provided n tends to infinity in such a way that h_r tends to zero $[r=0, 1, 2, \dots, (n-1).]$

This definition may also be stated in the form: Given the arbitrarily small positive number ε there is a positive number h such that

$$\left| \int_a^b F(x)\,dx - S_n \right| < \varepsilon, \text{ if } h_r < h, \ r = 0, 1, 2, \dots, (n-1). \quad (D_2)$$

It will often be convenient to call S_n an *approximation* to the integral I and equation (1) would read "$S_n = I$ approximately."

Cor. Instead of $F(\xi_r)$ we may take F_r where $m_r \leqq F_r \leqq M_r$.

Another form that is of frequent use in applications may be stated. In place of $F(\xi_r)$ put $F(x_r) + \alpha_r$ or $F(x_{r+1}) + \alpha_r$ where $|\alpha_r| < \varepsilon/(b-a)$ if $h_r < h$ so that α_r tends *uniformly* to zero when $h_r \to 0$. Let σ_n denote the sum

$$\sigma_n = \sum_{r=0}^{n-1} \{F(x_r) + \alpha_r\} h_r \quad \text{....................}(D_3)$$

$$= \sum_{r=0}^{n-1} F(x_r) h_r + \sum_{r=0}^{n-1} \alpha_r h_r.$$

Now if $h_r < h$ we have $|\Sigma \alpha_r h_r| < \{\varepsilon/(b-a)\}(\Sigma h_r)$, that is, $< \varepsilon$. Also, when n is sufficiently large $\Sigma F(x_r) h_r$ differs, by (D_2), from the integral I by less than ε. Hence $|I - \sigma_n| < 2\varepsilon$, when n is sufficiently large, say $n > N$, so that σ_n tends to I when n tends to infinity in such a way that the length of each sub-interval tends to zero.

108. Integrable Functions. The following classes of functions are integrable; the functions are supposed to be single-valued and bounded, and the range (a, b) of integration finite.

I. Continuous Functions. II. Monotonic Functions. III. Functions with Limited Variation.

I. *Continuous Functions.* The proof is given in § 106. A

constant C is a special case of a continuous function; but the integrability of a constant is obvious from the definition.

II. *Monotonic Functions.* Suppose first that the function $F(x)$ increases (or at least does not decrease) as x increases from a to b. If D is the division $[a, x_1, x_2, \dots, x_{n-1}, b]$ then $M_r = F(x_{r+1})$, $m_r = F(x_r)$ so that

$$S = F(x_1)(x_1 - a) + F(x_2)(x_2 - x_1) + \dots + F(b)(b - x_{n-1}),$$

$$s = F(a)(x_1 - a) + F(x_1)(x_2 - x_1) + \dots + F(x_{n-1})(b - x_{n-1}),$$

$$S - s = \sum_{r=0}^{n-1} \{F(x_{r+1}) - F(x_r)\}(x_{r+1} - x_r).$$

The differences $\{F(x_{r+1}) - F(x_r)\}$ are each positive or zero and their sum is $F(b) - F(a)$; therefore if each difference $(x_{r+1} - x_r)$ is less than h

$$S - s < h\Sigma\{F(x_{r+1}) - F(x_r)\} \text{ or } h\{F(b) - F(a)\}$$

so that $S - s < \varepsilon$ if $h < \varepsilon/\{F(b) - F(a)\}$, which is the condition of integrability.

If, next, $F(x)$ is a decreasing (non-increasing) function $M_r = F(x_r)$, $m_r = F(x_{r+1})$ and $S - s < \varepsilon$ if $h < \varepsilon/\{F(a) - F(b)\}$; in this case also the condition of integrability is satisfied.

Cor. A function that is bounded and has only a limited number of maxima and minima is integrable because the range of integration may be divided into a finite number of intervals in each of which the function is monotonic. (See § 109, Th. VII.)

III. *Functions with Limited Variation.* A function with limited variation can be expressed as the difference of two monotonic functions and therefore its integrability follows from II, if it be assumed, as will be proved immediately (§ 109), that the difference of two integrable functions is integrable.

109. General Theorems. The method of proof is simple. A division of the interval (a, b) is supposed to be made, the upper and lower sums S and s to be formed and the condition of integrability in one of its forms to be applied. In the application of the First Form it should be remembered that, by the Corollary to it, a function $F(x)$ is integrable over (a, b) provided there is *one* division of (a, b) for which $S - s < \varepsilon$ where ε

(as will always be assumed) denotes an arbitrarily small positive number.

It is supposed further that the interval of integration is (a, b) unless a different interval is expressly mentioned, so that the specification of the interval may be omitted in the enunciation of the Theorems.

THEOREM I. *If $F(x)$ is integrable so is $CF(x)$ where C is a constant and*

$$\int_a^b CF(x)\,dx = C\int_a^b F(x)\,dx.$$

If the division of (a, b) is such that $S - s$ for $F(x)$ is less than ε it is such that $S - s$ for $CF(x)$ is less than $|C|\,\varepsilon$ and since $|C|\,\varepsilon$ is, like ε, an arbitrarily small positive number $CF(x)$ is integrable. The equality of the integrals follows from the definition of an integral, for example, from the definition (D_1) of § 107.

THEOREM II. *If $F_1(x)$ and $F_2(x)$ are integrable so is their sum and their difference and*

$$\int_a^b \{F_1(x) \pm F_2(x)\}dx = \int_a^b F_1(x)\,dx \pm \int_a^b F_2(x)\,dx.$$

In the sub-interval (x_r, x_{r+1}) let M'_r, m'_r and M''_r, m''_r be the bounds of $F_1(x)$ and $F_2(x)$ respectively and S_1, s_1 and S_2, s_2 the respective sums. Let G_r, g_r and S_3, s_3 be the corresponding numbers for the sum $F_1(x) + F_2(x)$; then, as is easily seen,

$$G_r \leqq M'_r + M''_r,\ g_r \geqq m'_r + m''_r;\ G_r - g_r \leqq (M'_r - m'_r) + (M''_r - m''_r),$$

so that $$S_3 - s_3 \leqq (S_1 - s_1) + (S_2 - s_2).$$

It is easily proved that this relation also holds when G_r, g_r, S_3, s_3 are the corresponding numbers for $F_1(x) - F_2(x)$ because

$$G_r \leqq M'_r - m''_r,\ g_r \geqq m'_r - M''_r;\ G_r - g_r \leqq (M'_r - m''_r) - (m'_r - M''_r).$$

Hence if the division of (a, b) is such that $S_1 - s_1 < \frac{1}{2}\varepsilon$, $S_2 - s_2 < \frac{1}{2}\varepsilon$ it is such that $S_3 - s_3 < \varepsilon$ and therefore $F_1(x) \pm F_2(x)$ is integrable.

The relation between the integrals follows as before.

Cor. $$\int_a^b \left[\sum_{r=1}^m C_r F_r(x)\right] dx = \sum_{r=1}^m C_r \int_a^b F_r(x)\,dx$$

if $F_1(x), F_2(x), \ldots, F_m(x)$ are integrable and $C_1, C_2, \ldots, C_m$ constants, m being a finite integer.

THEOREM III. *If $F_1(x)$ and $F_2(x)$ are integrable so is their product.*

First, suppose that $F_1(x)$ and $F_2(x)$ are both positive and use the same notation as in the proof of Theorem II, the numbers G_r, g_r, S_3, s_3 referring to the product $F_1(x)F_2(x)$. A little consideration shows that $G_r \leqq M_r'M_r''$ and $g_r \geqq m_r'm_r''$; therefore

$$G_r - g_r \leqq M_r'M_r'' - m_r'm_r'' = M_r'(M_r'' - m_r'') + m_r''(M_r' - m_r')$$

so that $$G_r - g_r < A(M_r'' - m_r'') + B(M_r' - m_r'),$$

and $$S_3 - s_3 < A(S_2 - s_2) + B(S_1 - s_1),$$

where A and B are upper bounds of $F_1(x)$ and $F_2(x)$ in (a, b).

Hence if the division of (a, b) is such that $S_1 - s_1 < \varepsilon$ and $S_2 - s_2 < \varepsilon$ it is such that $S_3 - s_3 < (A + B)\varepsilon$ and therefore is arbitrarily small, so that the product $F_1(x)F_2(x)$ is integrable.

Next, if $F_1(x)$ and $F_2(x)$ are not both positive for $a \leqq x \leqq b$ there are positive constants C_1 and C_2 such that $F_1(x) + C_1$ and $F_2(x) + C_2$ are both positive in (a, b) and therefore the product $(F_1 + C_1)(F_2 + C_2)$ integrable. But

$$F_1F_2 = (F_1 + C_1)(F_2 + C_2) - C_2F_1 - C_1F_2 - C_1C_2$$

and therefore F_1F_2 is integrable since it is the sum of integrable functions.

Cor. If each of the m functions $F_1(x)$, $F_2(x)$, ..., $F_m(x)$ is integrable so is their product, m being a finite integer.

THEOREM IV. *If $F(x)$ is integrable in (a, b) so is $1/F(x)$ provided $|F(x)| > c > 0$ for $a \leqq x \leqq b$ where c is a constant.*

In the sub-interval (x_r, x_{r+1}) let M_r, m_r and G_r, g_r be the upper and lower bounds of $F(x)$ and $1/F(x)$ respectively.

First, suppose that $F(x)$ is either always positive, $F(x) > c$, or else always negative, $F(x) < -c$, for $a \leqq x \leqq b$. In this case $G_r = 1/m_r$ and $g_r = 1/M_r$ while the product M_rm_r is positive and greater than c^2 Therefore

$$G_r - g_r = \frac{1}{m_r} - \frac{1}{M_r} = \frac{M_r - m_r}{M_rm_r} < \frac{1}{c^2}(M_r - m_r).$$

Next, suppose that $F(x)$ takes both positive and negative values in (x_r, x_{r+1}) so that $M_r > 0$ and $m_r < 0$. Let m_r' be the lower bound of the positive values of $F(x)$ and M_r' the upper bound of the negative values of $F(x)$ in (x_r, x_{r+1}); in this case $G_r = 1/m_r'$ and $g_r = 1/M_r'$. Now M_r' and m_r are both negative

and $(-M'_r) \leqq (-m_r)$ while $(-M'_r) > c$; also $M_r \geqq m'_r$. Hence

$$G_r - g_r = \frac{1}{m'_r} - \frac{1}{M'_r} = \frac{m'_r - M'_r}{(-M'_r)m'_r} < \frac{1}{c^2}(M_r - m_r),$$

and therefore if S, s and S_1, s_1 are the sums for $F(x)$ and $1/F(x)$ respectively $S_1 - s_1 < (S-s)/c^2$ so that $1/F(x)$ is integrable.

Cor. If $F(x)$ satisfies the conditions of the Theorem and if $F_1(x)$ is integrable so is $F_1(x)/F(x)$. This is now simply a particular case of Theorem III.

THEOREM V. *Any rational function $\varphi(F_1, F_2, \ldots, F_m)$ of m integrable functions $F_1(x)$, $F_2(x)$, ..., $F_m(x)$, m being a finite integer, is integrable provided the lower bound of $|\varphi(F_1, F_2, \ldots, F_m)|$ is positive (not zero).*

This follows at once from the preceding four Theorems.

THEOREM VI. *If $F(x)$ is integrable so is $|F(x)|$ but $|F(x)|$ may be integrable and $F(x)$ not integrable. Further*

$$\left|\int_a^b F(x)\,dx\right| \leqq \int_a^b \left|\, F(x)\,\right| dx.$$

If y, z are any two numbers and $|y| = \eta$, $|z| = \zeta$ then

$$|y - z| \geqq |\eta - \zeta|$$

so that the oscillation of $|F(x)|$ in any sub-interval cannot exceed that of $F(x)$; hence $|F(x)|$ is integrable if $F(x)$ is. The relation between the two integrals follows at once from the form (D_1), § 107, of the definition of an integral.

That $|F(x)|$ may be integrable but $F(x)$ not integrable may be seen by considering the (somewhat artificial) function $F(x)$, defined for the interval (0, 1) as follows: $F(x) = 1$ for irrational values but $F(x) = -1$ for rational values of x in (0, 1). In this case $|F(x)| = 1$ and is therefore integrable. On the other hand, in any sub-interval the upper and lower bounds of $F(x)$ are 1 and -1 respectively so that $S - s = 2$ whatever division of (0, 1) be made and therefore $F(x)$ is not integrable.

THEOREM VII. *If the interval (a, b) is divided into m partial intervals (a, a_1), (a_1, a_2), ..., (a_{m-1}, b) by the fixed numbers $a_1, a_2, \ldots, a_{m-1}$ where $a < a_1 < a_2 < \ldots < a_{m-1} < b$ and if $F(x)$ is integrable over (a, b) it is integrable over each partial interval (a, a_1), (a_1, a_2), ..., (a_{m-1}, b). Conversely, if $F(x)$ is integrable over each partial interval it is integrable over the whole interval (a, b). In both cases*

$$\int_a^b F(x)\,dx = \int_a^{a_1} F(x)\,dx + \int_{a_1}^{a_2} F(x)\,dx + \ldots + \int_{a_{m-1}}^b F(x)\,dx.$$

The proof is obvious. If there is a division of (a, b), the numbers $a_1, a_2, \dots, a_{m-1}$ being fixed points of the division, such that $S-s$ for the whole range is less than ε, then $S-s$ for any one of the partial intervals is certainly less than ε. Again, if there is a division of (a, b) such that $S-s$ for each partial interval is less than ε/m then $S-s$ for the whole interval (a, b) is less than $m(\varepsilon/m)$ or ε.

110. Discontinuities. It has been seen in § 108 that every continuous function is integrable; the third form of the condition of integrability (§ 106) shows, however, that a bounded function may be discontinuous and yet integrable. The following theorem throws some light on what may be called "admissible discontinuities."

THEOREM I. *A bounded function $F(x)$ is integrable over (a, b)* (i) *if there is only a finite number, m say, of points of discontinuity in (a, b), and* (ii) *if there is an infinite number of points of discontinuity in (a, b) provided this infinite set of points has only a finite number of limiting points in (a, b).*

First, let there be only one point of discontinuity, c, and let $|F(x)|$ be less than K for every value of x in (a, b). Choose δ ($\delta>0$) so that the length 2δ of the sub-interval $(c-\delta, c+\delta)$ may be less than $\varepsilon/4K$; then the part of $S-s$ arising from the interval $(c-\delta, c+\delta)$ can not exceed the product of $2K$ and 2δ (the length of the sub-interval), that is, cannot exceed $\frac{1}{2}\varepsilon$. In the intervals $(a, c-\delta)$ and $(c+\delta, b)$ the function $F(x)$ is continuous so that there is a division of the intervals $(a, c-\delta)$ and $(c+\delta, b)$ such that the part of $S-s$ arising from these two intervals jointly is less than $\frac{1}{2}\varepsilon$. Thus there is a division of the whole interval (a, b) for which $S-s$ is less than ε and therefore $F(x)$ is integrable over (a, b).

Next, let there be m points of discontinuity $c_1, c_2, \dots, c_m$ and enclose these in sub-intervals $(c_r-\delta_r, c_r+\delta_r)$, $r=1, 2, \dots, m$, such that the sum $(2\Sigma\delta_r)$ of their lengths is less than $\varepsilon/4K$, where K has the same meaning as in the first case. The part of $S-s$ arising from these m sub-intervals cannot exceed $2K\times(2\Sigma\delta_r)$, that is, $\frac{1}{2}\varepsilon$; on the other hand, in the partial intervals $(a, c_1-\delta_1)$, $(c_1+\delta_1, c_2-\delta_2)$, $\dots$, $(c_m+\delta_m, b)$ the function $F(x)$ is continuous and therefore the part of $S-s$ arising from

these partial intervals jointly can by a suitable division of the intervals be made less than $\frac{1}{2}\varepsilon$. Hence there is a division of (a, b) for which $S-s$ is less than ε and therefore $F(x)$ is integrable over (a, b).

Finally, suppose that the set of points c_r for which $F(x)$ is discontinuous is infinite but has only a finite number, μ say, of limiting points. If there is *only one* limiting point—say ξ—all but a finite number of the points $c_1, c_2, \ldots$ can be enclosed in a sub-interval $(\xi-\delta, \xi+\delta)$ where $2\delta < \varepsilon/8K$ (K as before); the remaining points $c_1, c_2, \ldots$ can be enclosed in sub-intervals whose total length is less than $\varepsilon/8K$ and there is left a finite number of partial intervals in each of which $F(x)$ is continuous. The contribution to $S-s$ from the interval $(\xi-\delta, \xi+\delta)$ is less than $2K\times(\varepsilon/8K)$ or $\frac{1}{4}\varepsilon$ and the contribution from the sub-intervals that enclose $c_1, c_2, \ldots$ is also less than $\frac{1}{4}\varepsilon$; further, there is a division of the partial intervals in which $F(x)$ is continuous for which the contribution to $S-s$ is less than $\frac{1}{2}\varepsilon$. Hence on the whole there is a division of (a, b) for which $S-s$ is less than ε so that $F(x)$ is integrable over (a, b).

In the same way the proof is carried out when there are μ limiting points.

The theorem just proved leads to an interesting result. If the *bounded* functions $F(x)$ and $f(x)$ are equal for the range $a \leqq x \leqq b$, except for the values $c_1, c_2, \ldots, c_m$ of x, and if $F(x)$ is integrable over (a, b) so is $f(x)$, and further

$$\int_a^b f(x)\,dx = \int_a^b F(x)\,dx.$$

Suppose that $|F(x)| < H$ and $|f(x)| < K$ when $a \leqq x \leqq b$. Let the points c_r be enclosed in sub-intervals $(c_r-\delta_r, c_r+\delta_r)$; then the values of $S-s$ for $F(x)$ and $f(x)$ respectively differ only in the parts that arise from the m sub-intervals $(c_r-\delta_r, c_r+\delta_r)$. But that difference cannot exceed $(2H+2K)\times(2\Sigma\delta_r)$ and will therefore be less than ε if $\Sigma\delta_r$ is chosen (as is possible) to be less than $\varepsilon/(4H+4K)$. Since $F(x)$ is integrable over (a, b) there is a division of (a, b) such that $S-s$ for $F(x)$ is less than ε; therefore for that division and for the function $f(x)$ the difference $S-s$ is less than 2ε so that $f(x)$ is integrable over (a, b). That the two integrals are equal follows from the facts (i) that the integrals are constants and (ii) that their difference depends

solely on the contribution from the sub-intervals $(c_r - \delta_r, c_r + \delta_r)$, which can be made arbitrarily small; but two constants whose difference is arbitrarily small are equal. Hence we find the following theorem:

THEOREM II. *The values of an integrable bounded function $f(x)$ may be arbitrarily changed at any finite number of points in the range of integration (a, b) without changing the value of the integral over (a, b) provided the new values of $F(x)$ are finite.*

It is easy to see that this theorem is still true when the number of points c_r for which the value of $F(x)$ is changed (the new values of $F(x)$ being finite) is infinite provided the set (c_r) of points has only a finite number of limiting points.

Ex. 1. $F(x)$ is defined for the interval $(0, 1)$ by the condition that if r is a positive integer, $r = 1, 2, 3, \ldots$

$$F(x) = 2rx \text{ when } \frac{1}{r+1} < x < \frac{1}{r};$$

prove that $F(x)$ is integrable over $(0, 1)$. (Nielsen, *Elemente der Funktionenlehre*, p. 143.)

If δ is positive and sufficiently small,

$$F\left(\frac{1}{r} - \delta\right) = 2r\left(\frac{1}{r} - \delta\right),\ r = 1, 2, 3, \ldots;$$

$$F\left(\frac{1}{r} + \delta\right) = 2(r-1)\left(\frac{1}{r} + \delta\right),\ r = 2, 3, \ldots$$

so that $F(r^{-1} - \delta) \to 2$ and $F(r^{-1} + \delta) \to 2(1 - r^{-1})$ when $\delta \to 0$. Hence, however $F(x)$ may be defined for $x = 1/r$, $r = 2, 3, \ldots$ the points $x = 1/r$ are points of discontinuity. $F(x)$ will be continuous for x tending to 1 if $F(1) = 2$ and 2 will be taken as the value of $F(1)$.

The set of numbers $\frac{1}{2}, \frac{1}{3} \ldots, \frac{1}{r}, \ldots$ has only one limiting point, namely the point $x = 0$; let $F(x) = 2$ when $x = 0$.

Now enclose the point $1/r$ in the sub-interval $(r^{-1} - \delta_r, r^{-1} + \delta_r)$ where $\delta_r = \varepsilon/2^{r+1}$, $r = 2, 3, \ldots, (m+1)$; further enclose the limiting point 0 in the sub-interval $(0, p^{-1})$ where $m+1 < p < m+2$. The total length of these sub-intervals is

$$\frac{1}{p} + \varepsilon \sum_{r=2}^{m+1} \frac{1}{2^r} = \frac{1}{p} + \frac{1}{2}\varepsilon\left(1 - \frac{1}{2^m}\right) < \frac{1}{p} + \frac{1}{2}\varepsilon,$$

and is less than ε if m is chosen so that $m + 1 > 2/\varepsilon$ and therefore $1/p < \varepsilon/2$. Hence $F(x)$ is integrable over $(0, 1)$.

The integrability of $F(x)$ follows at once from the fact that the set of points for which $F(x)$ is discontinuous has only one limiting point;

the above determination of the sub-intervals that enclose the points of discontinuity is merely made in order to indicate one way of securing the required sub-intervals.

For the evaluation of the integral see § 112.

Ex. 2. Show that $\sin(1/x)$ is integrable over any finite range, whether the range includes the point $x=0$ or not.

Ex. 3. If $F(x)$ is bounded and monotonic for the range $a \leqq x \leqq b$, prove that the limits for δ tending to zero of $F(c-\delta)$ and $F(c+\delta)$, where $a<c<b$, both exist; also that, if $\delta>0$, the limits of $F(a+\delta)$ and $F(b-\delta)$ both exist.

111. Properties of the Integral. It has been assumed up to this stage that the upper limit b is greater than the lower limit a; this restriction will now be removed.

If $b=a$ the integral is defined to be zero.

Definition 1
$$\int_a^a F(x)\,dx=0.$$

If $b<a$ the numbers in the division $[a, x_1, x_2, \ldots x_{n-1}, b]$ of the interval (a, b) satisfy the relations
$$a>x_1>x_2>\ldots>x_{n-1}>b,$$
and each difference $x_{r+1}-x_r$ is negative. The sums S, s and their limits L, l simply have their signs changed; hence the definition:

Definition 2.
$$\int_a^b F(x)\,dx=-\int_b^a F(x)\,dx.$$

If the three numbers a, b, c all lie within an interval over which $F(x)$ is integrable we have
$$\int_a^c F(x)\,dx+\int_c^b F(x)\,dx+\int_b^a F(x)\,dx=0,$$
as an equivalent form of the equation
$$\int_a^c F(x)\,dx+\int_c^b F(x)\,dx=\int_a^b F(x)\,dx,$$
which was previously (§ 109, Th. VII) proved for the relation $a<c<b$.

In § 124, pp. 298-301 of the *Elementary Treatise*, some inequalities between integrals are proved, but these all depend on Theorem III, p. 298; when that theorem has been proved for the integral of a bounded function Theorems V, VI and VII

of that article will then hold for the integrals of bounded functions. We now prove that theorem.

THEOREM. *If* $a<b$ *and* $F(x) \geqq 0$, $\int_a^b F(x)\,dx \geqq 0$; *if* $F(x) \leqq 0$, $\int_a^b F(x)\,dx \leqq 0$.

If $F(x) \geqq 0$ the lower sum s cannot be negative and therefore the integral cannot be negative; similarly, if $F(x) \leqq 0$ the integral cannot be positive, since the upper sum S cannot be positive.

On account of their frequent use the two **Mean Value Theorems** for bounded integrable functions are stated:

First Theorem of Mean Value. *If* $a<b$, $\varphi(x) \geqq 0$, $g \leqq \psi(x) \leqq G$ *for* $a \leqq x \leqq b$, *then*

$$\text{(i)}\quad g\int_a^b \varphi(x)\,dx \leqq \int_a^b \varphi(x)\psi(x)\,dx \leqq G\int_a^b \varphi(x)\,dx\,;$$

$$\text{(ii)}\quad \int_a^b \varphi(x)\psi(x)\,dx = K\int_a^b \varphi(x)\,dx,\ g \leqq K \leqq G\,;$$

if $\psi(x)$ *is not merely integrable but continuous for* $a \leqq x \leqq b$,

$$\text{(iii)}\quad \int_a^b \varphi(x)\psi(x)\,dx = \psi\{a+\theta(b-a)\}\int_a^b \varphi(x)\,dx,\ 0<\theta<1.$$

Equations (ii) and (iii) are valid if $a>b$.

Second Theorem of Mean Value. *If for* $a \leqq x \leqq b$ *the function* $\varphi(x)$ *is bounded, positive and decreases (or at least does not increase) as* x *increases, and if* $\psi(x)$ *is bounded and integrable, then*

$$\text{(i)}\quad \int_a^b \varphi(x)\psi(x)\,dx = \varphi(a+0)\int_a^\xi \psi(x)\,dx,\ a \leqq \xi \leqq b\,;$$

if $\varphi(x)$ *is simply bounded and monotonic, then*

$$\text{(ii)}\quad \int_a^b \varphi(x)\psi(x)\,dx = \varphi(a+0)\int_a^\xi \psi(x)\,dx + \varphi(b-0)\int_\xi^b \psi(x)\,dx,\ a \leqq \xi \leqq b.$$

The proof already given of the Second Theorem (*E.T.* pp. 452-454) is valid for the theorem as now stated. The following points should be noted:

(1) $\varphi(x)$, being bounded and monotonic, is integrable and the product of the two integrable functions $\varphi(x)$ and $\psi(x)$ is integrable;

(2) the limits $\varphi(a+0)$ and $\varphi(b-0)$ exist, and if a, b are points

of discontinuity of $\varphi(x)$ these are to be taken as the values of $\varphi(a)$ and $\varphi(b)$; further, if $a<x_r<b$, and if x_r is a point of discontinuity of $\varphi(x)$, the value of $\varphi(x_r)$ may be taken to be $\varphi(x_r-0)$ or $\varphi(x_r+0)$ or any number between these;

(3) as will be proved in the next article, the integral $\int_a^x \psi(t)\,dt$ is a *continuous* function of x and therefore there is a value ξ of x such that, for $x=\xi$, that integral is equal to the mean value M (*E.T.* p. 452);

(4) by § 109, Theorem VI, $\left|\int_{x_{r-1}}^{x_r} \psi(x)\,dx\right| \leqq \int_{x_{r-1}}^{x_r} |\psi(x)|\,dx$

so that if $|\psi(x)|<K$ for $a \leqq x \leqq b$, and if n is chosen so large that each difference $(x_r - x_{r-1})$ is less than ε/K, the integral just written will be less than ε;

(5) $\sum_{r=1}^{n} \{\varphi(x_{r-1}) - \varphi(x_r)\} \leqq \varphi(a+0) - \varphi(b-0)$.

It will be a good exercise to go carefully through the proof.

It is sometimes more convenient to express the theorem, not in terms of the mean value M but in terms of the two numbers between which M lies. Let $f(x)$ be the integral from which the mean value M is derived, namely,

$$f(x) = \int_a^x \psi(t)\,dt.$$

As x varies from a to b, the function $f(x)$ being continuous will take once at least every value between its lower bound, g say, which is in this case the least value of $f(x)$, and its upper bound or greatest value, G say. The Mean Value Theorem may therefore be expressed in a third form, namely,

(iii) $g\varphi(a+0) \leqq \int_a^b \varphi(x)\,\psi(x)\,dx \leqq G\varphi(a+0)$.

Cor. In form (i) let $x=a+b-u$, $\varphi(x)=\varphi_1(u)$, $\psi(x)=\psi_1(u)$; $\varphi_1(u)$ is bounded, positive and increases (or at least does not decrease) as u increases from a to b. Thus

$$\int_a^b \varphi_1(u)\psi_1(u)\,du = \varphi_1(b-0)\int_\eta^b \psi_1(u)\,du, \quad a \leqq \eta \leqq b,$$

or

$$g_1\varphi_1(b-0) \leqq \int_a^b \varphi_1(u)\psi_1(u)\,du \leqq G_1\varphi_1(b_1-0),$$

where g_1, G_1 are the least and greatest values of $\int_u^b \psi_1(u)\,du$ as u varies from a to b.

The change of variable is valid by § 114. Here

$$\varphi(a+0)=\varphi_1(b-0)$$

and $$\int_a^\xi \psi(x)\,dx=\int_{a+b-\xi}^b \psi_1(u)\,du=\int_\eta^b \psi_1(u)\,du,\quad a+b-\xi=\eta.$$

See the other forms of the Second Theorem of Mean Value stated in § 112, Ex. 2.

112. The Integral as a Function of its Limits. Take x as the upper limit of the integral and, to avoid ambiguity, t as the variable of integration; $F(x)$ is supposed to be bounded and integrable for $a \leqq x \leqq b$.

Let $$\varphi(x)=\int_a^x F(t)\,dt. \qquad\qquad (1)$$

and for x put $x+h$, $a \leqq x+h \leqq b$; h may be either positive or negative. Now

$$\varphi(x+h)=\int_a^{x+h} F(t)\,dt=\int_a^x F(t)\,dt+\int_x^{x+h} F(t)\,dt,$$

so that $$\varphi(x+h)-\varphi(x)=\int_x^{x+h} F(t)\,dt. \qquad\qquad (2)$$

The function $F(t)$ is bounded, say $|F(t)|<K$; therefore

$$|\varphi(x+h)-\varphi(x)|<K\,|h|,$$

so that $\varphi(x+h)\to\varphi(x)$ when $h\to 0$. We thus find the very important theorem:

THEOREM I. *The integral of a bounded function $F(x)$ is continuous whether $F(x)$ is continuous or not.*

Again, if $F(t)$ is continuous for $x-|h| \leqq t \leqq x+|h|$, the First Theorem of Mean Value gives ($h>0$ or $h<0$)

$$\int_x^{x+h} F(t)\,dt=hF(x+\theta h),\ 0<\theta<1,$$

and therefore

$$\varphi'(x)=\mathop{\mathcal{L}}_{h\to 0}\frac{\varphi(x+h)-\varphi(x)}{h}=\mathop{\mathcal{L}}_{h\to 0}F(x+\theta h)=F(x) \qquad (3)$$

that is $$\frac{d}{dx}\int_a^x F(t)\,dt=F(x). \qquad\qquad (3')$$

Cor. $$\frac{d}{dx}\int_x^b F(t)\,dt=-\frac{d}{dx}\int_b^x F(t)\,dt=-F(x). \qquad\qquad (3'')$$

From the equation (3) the following fundamental theorem is deduced :-

THEOREM II. *If $F(x)$ is continuous for $a \leqq x \leqq b$ and if $F(x) = f'(x)$, then*

$$\int_a^b F(x)\,dx = f(b) - f(a). \qquad \text{.....................(4)}$$

For, if $\varphi(x)$ is defined by (1) it has been proved that $\varphi(x)$ is continuous and that $\varphi'(x) = F(x)$; hence $\varphi'(x) - f'(x)$ is zero and therefore (§ 34) $\varphi(x) - f(x)$ is a constant, C say. Thus

$$\int_a^x F(t)\,dt - f(x) = C, \text{ or, } \int_a^x F(t)\,dt = f(x) + C \text{ ....(5)}$$

When $x = a$, the integral is zero, so that $C = -f(a)$, and therefore

$$\int_a^x F(t)\,dt = f(x) - f(a). \qquad \text{.....................(5')}$$

If $x = b$ the equation (5′) has the same meaning as equation (4), since the variable of integration may be taken to be x instead of t.

When the integral in (5) is considered simply as a function of its upper limit x we may omit the lower limit a and write

$$\int^x F(t)\,dt = f(x) + \text{const., or, } \int F(x)\,dx = f(x) + \text{const.} \quad \text{...(5'')}$$

The two symbols $\int^x F(t)\,dt$ and $\int F(x)\,dx$ mean the same thing, namely "the indefinite integral of $F(x)$ with respect to x"; since $f'(x) = F(x)$ we thus verify the usual rule that "the derivative of the integral is equal to the integrand."

When $F(x)$ is continuous the integral $\varphi(x)$ exists, and when the function $f(x)$ has been found it may be said that the integral has been "evaluated" and the equations (5), (5′), (5″) give the "value" of the integral. It has to be noted, however, that Theorem II has been proved on the assumption that $F(x)$ is continuous for the *closed* range (a, b) or (a, x) if $a < x < b$. When $F(x)$ is not continuous the theorem given in (3) requires modification and therefore also Theorem II.

Discontinuity of $F(x)$. Suppose that $F(x)$ is continuous in (a, b) except for the one value c of x, and that the discontinuity is of the *first kind*. If $a < c < b$ the limits $F(c - 0)$ and $F(c + 0)$ exist but are not equal; if c is a or b the limit $F(a + 0)$ or $F(b - 0)$ exists but is not equal to $F(a)$ or $F(b)$.

In equation (2) let $x=c$ and take $h>0$. Since $\varphi(x)$ is continuous $\varphi(c)$ is a definite number and the Mean Value Theorem gives

$$\varphi(c+h)-\varphi(c)=hF(c+\theta h),$$

so that
$$\varphi'(c+0)=\underset{h\to 0}{\mathcal{L}}\frac{\varphi(c+h)-\varphi(c)}{h}=F(c+0);$$

in the same way it is seen that $\varphi'(c-0)=F(c-0)$.

On the other hand, if c is a point of discontinuity of the *second kind* either $F(c-0)$ or $F(c+0)$ or both will not be definite and one or both of the derivatives $\varphi'(c-0)$, $\varphi'(c+0)$ will not exist. Thus when c is a point of discontinuity for $F(x)$ it is also a point of discontinuity of the same kind for $\varphi'(x)$ while, it must always be remembered, it is a point of continuity for $\varphi(x)$.

When $F(x)$ is integrable over (a, b) and is discontinuous, say, for x equal to $c_1, c_2, \dots, c_m$ where $a \leqq c_1$ and $b \geqq c_m$, enclose c_r in the sub-interval $(c_r-\delta_r, c_r+\delta_r)$; we may put each δ_r equal to δ where δ is positive and so small that when x is in the partial interval (c_r, c_{r+1}) we shall have $c_r+\delta \leqq x \leqq c_{r+1}-\delta$. The limit for $\delta \to 0$ of the sum

$$\int_a^{c_1-\delta}F(x)\,dx+\sum_{r=1}^{m-1}\int_{c_r+\delta}^{c_{r+1}-\delta}F(x)\,dx+\int_{c_m+\delta}^{b}F(x)\,dx$$

is, since each integral is continuous,

$$\int_a^{c_1}F(x)\,dx+\sum_{r=1}^{m-1}\int_{c_r}^{c_{r+1}}F(x)\,dx+\int_{c_m}^{b}F(x)\,dx=\int_a^bF(x)\,dx.$$

Now $F(x)$ is continuous for $c_r+\delta \leqq x \leqq c_{r+1}-\delta$, and therefore, by Theorem II, if $F(x)=f_r'(x)$ for $c_r+\delta \leqq x \leqq c_{r+1}-\delta$ we have

$$\int_{c_r}^{c_{r+1}}F(x)\,dx=\underset{\delta\to 0}{\mathcal{L}}\,[f_r(c_{r+1}-\delta)-f_r(c_r+\delta)]=f_r(c_{r+1})-f_r(c_r).$$

Hence, for symmetry, denoting a by c_0 and b by c_{m+1}, we find

$$\int_a^bF(x)\,dx=\sum_{r=0}^{m}\,[f_r(c_{r+1})-f_r(c_r)].$$

In practice, when it is known that the integral exists, we may at once write

$$\int_{c_r}^{c_{r+1}}F(x)\,dx=f_r(c_{r+1})-f_r(c_r).$$

Ex. 1. As an illustration, take Ex. 1 of § 110. In this case

$$F(x) = 2rx \text{ when } 1/(r+1) < x < 1/r.$$

Let $2rx = f_r'(x)$ and therefore

$$\int_{\frac{1}{r+1}}^{\frac{1}{r}} F(x)dx = \int_{\frac{1}{r+1}}^{\frac{1}{r}} 2rxdx = r\left\{\frac{1}{r^2} - \frac{1}{(r+1)^2}\right\}.$$

The integral of $F(x)$ is obtained by giving to r the values 1, 2, ... m, adding the partial integrals and letting m tend to infinity ($1/p$ tends to zero when m tends to infinity). Hence

$$\int_{\frac{1}{m+1}}^{1} F(x)dx = \sum_{r=1}^{m} r\left\{\frac{1}{r^2} - \frac{1}{(r+1)^2}\right\} = \sum_{r=1}^{m} \frac{1}{r^2} - \frac{m}{(m+1)^2}$$

and

$$\int_0^1 F(x)dx = \sum_{r=1}^{\infty} \frac{1}{r^2} = \frac{\pi^2}{6}.$$

Ex. 2. Prove that in the form (i) of the Second Theorem of Mean Value (§ 111) it is admissible to substitute A in place of $\varphi(a+0)$ provided that A is greater than $\varphi(a+0)$, A being finite.

The value of the integral of $\varphi(x)\psi(x)$ over the range (a, b) is not changed by substituting $A\psi(a)$ in place of $\varphi(a+0)\psi(a)$, by Theorem II of § 110; further, the monotonic character of $\varphi(x)$ is preserved since $A > \varphi(a+0)$ so that the proof is still valid.

Similarly, in the form (ii) we may put A in place of $\varphi(a+0)$ and B in place of $\varphi(b-0)$ provided the monotonic character of $\varphi(x)$ is preserved; that is, $A > \varphi(a+0) > \varphi(x) > \varphi(b-0) > B$ if $\varphi(x)$ decreases or $A < \varphi(a+0) < \varphi(x) < \varphi(b-0) < B$ if $\varphi(x)$ increases. Hence the two forms

$$\text{(i}a\text{)} \quad \int_a^b \varphi(x)\psi(x)dx = A\int_a^\xi \psi(x)dx, \quad A > \varphi(a+0) > \varphi(x),$$

$$\text{(ii}a\text{)} \quad \int_a^b \varphi(x)\psi(x)dx = A\int_a^\xi \psi(x)dx + B\int_\xi^b \psi(x)dx,$$

where in (ii$a$) A and B suit the monotonic character of $\varphi(x)$ as explained above.

113. Examples. One method of dividing the interval (a, b) into n sub-intervals is to make each sub-interval of the same length h, where $h = (b-a)/n$. In this case the integral of $F(x)$ over (a, b) is the limit for $h \to 0$ of the sum

$$\text{(i)} \quad \sum_{r=1}^{n} F(\xi_r)h, \quad a + (r-1)h \leq \xi_r \leq a + rh;$$

if we first put $\xi_r = a + rh$ and then $\xi_r = a + (r-1)h$ we find the following two sums, which are specially useful:

$$\text{(ii)} \quad \sum_{r=1}^{n} F(a+rh).h, \qquad \text{(iii)} \quad \sum_{r=1}^{n} F\{a+(r-1)h\}.h.$$

When $h \to 0$ each of the sums (i), (ii) and (iii) tends to

$$\int_a^b F(x)\,dx.$$

Ex. 1. If $S_n = \frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n}$, prove that $S_n \to \log 2$ when $n \to \infty$.

Here we may write

$$S_n = \sum_{r=1}^{n} \frac{1}{1+rh} \cdot h,\ h = \frac{1}{n},$$

and comparison with form (ii) above shows that we may take $F(x) = 1/x$, $a = 1$, $b = 2$ or $F(x) = 1/(1+x)$, $a = 0$, $b = 1$. Therefore the limit of S_n is

$$\int_1^2 \frac{dx}{x} = \int_0^1 \frac{dx}{1+x} = \log 2.$$

Cor. It is easy to prove that $S_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{1}{2n-1} - \frac{1}{2n}$ and thus to deduce that $\log 2 = \sum_1^{\infty} (-1)^{n-1}/n$.

Ex. 2. If $F(x)$ and $F'(x)$ are continuous for $a \leqq x \leqq b$, and if

$$S_n = \sum_{r=1}^{n} F(a+rh)h,\ I = \int_a^b F(x)\,dx$$

where $h = (b-a)/n$, prove that the limit of $n(S_n - I)$ for $n \to \infty$ is

$$\tfrac{1}{2}(b-a)\{F(b) - F(a)\}.$$

For brevity, let $a + rh = x_r$, $a = x_0$, $b = x_n$; then

$$S_n = \sum_{r=1}^{n} F(x_r)\int_{x_{r-1}}^{x_r} dx, \text{ since } \int_{x_{r-1}}^{x_r} dx = x_r - x_{r-1} = h,$$

$$I = \sum_{r=1}^{n} \int_{x_{r-1}}^{x_r} F(x)\,dx,\ S_n - I = \sum_{r=1}^{n} \int_{x_{r-1}}^{x_r} [F(x_r) - F(x)]\,dx.$$

But if $x_{r-1} \leqq x \leqq x_r$, $F(x) = F(x_r) - (x_r - x)F'(\xi_r)$, $x_{r-1} \leqq \xi_r \leqq x_r$. Therefore

$$S_n - I = \sum_{r=1}^{n} \int_{x_{r-1}}^{x_r} (x_r - x)\,F'(\xi_r)\,dx. \quad \dots\dots\dots\dots(1)$$

Now suppose that $g_r \leqq F'(x) \leqq G_r$ for $x_{r-1} \leqq x \leqq x_r$; then,

$$g_r \int_{x_{r-1}}^{x_r} (x_r - x)\,dx \leqq \int_{x_{r-1}}^{x_r} (x_r - x)\,F'(\xi_r)\,dx \leqq G_r \int_{x_{r-1}}^{x_r} (x_r - x)\,dx,$$

that is,
$$\tfrac{1}{2}h^2 g_r \leqq \int_{x_{r-1}}^{x_r} (x_r - x)\,F'(\xi_r)\,dx \leqq \tfrac{1}{2}h^2 G_r.$$

Hence, multiplying both sides of equation (1) by n and noting that $nh = b - a$, we find

$$\tfrac{1}{2}(b-a)\sum_{r=1}^{n} g_r h \leqq n(S_n - I) \leqq \tfrac{1}{2}(b-a)\sum_{r=1}^{n} G_r h.$$

But by the form (i) above, when $h \to 0$, both $\Sigma g_r h$ and $\Sigma G_r h$ tend to

$$\int_a^b F'(x)dx = F(b) - F(a),$$

and therefore

$$\mathop{\mathcal{L}}_{n\to\infty} n(S_n - I) = \tfrac{1}{2}(b-a)\{F(b) - F(a)\}.$$

Ex. 3. If I and h have the same meaning as in Example 2, but if now $F''(x)$ is also continuous in (a, b) and

$$S_n = \sum_{r=1}^{n} F\left(a + \frac{2r-1}{2}h\right)h,$$

show that the limit of $n^2(I - S_n)$ for $n \to \infty$ is

$$\tfrac{1}{24}(b-a)^2\{F'(b) - F'(a)\}.$$

Proceed exactly as in Example 2; note that for the interval (x_{r-1}, x_r), if $a + \dfrac{2r-1}{2}h = c$, we have $F(x) - F(c) = (x-c)F'(c) + \frac{1}{2}(x-c)^2 F''(\xi_r)$ and therefore

$$\int_{x_{r-1}}^{x_r} \left\{F(x) - F\left(a + \frac{2r-1}{2}h\right)\right\}dx = \tfrac{1}{2}\int_{x_{r-1}}^{x_r}\left(x - a - \frac{2r-1}{2}h\right)^2 F''(\xi_r)dx,$$

where $x_{r-1} \leqq \xi_r \leqq x_r$, and this integral lies between $\frac{1}{24}g_r h^3$ and $\frac{1}{24}G_r h^3$, G_r and g_r being the upper and lower bounds of $F''(x)$ in the interval (x_{r-1}, x_r). Hence

$$\mathop{\mathcal{L}}_{n\to\infty} n^2(I - S_n) = \frac{(b-a)^2}{24}\int_a^b F''(x)dx = \frac{(b-a)^2}{24}\{F'(b) - F'(a)\}.$$

Ex. 4. Show that $\mathop{\mathcal{L}}\limits_{n\to\infty} n(\log 2 - S_n) = +\tfrac{1}{4}$ where S_n has the same meaning as in *Example* 1.

Apply *Example* 2; $a = 1$, $b = 2$, $F(x) = 1/x$.

114. Transformations of the Integral. No change is required in the proof of the formula of *Integration by Parts* (*E.T.* p. 282) when the functions that appear in the formula are continuous; the use of the formula is in practice confined to this case. The formula for *Change of Variable*, however, requires a new proof.

Suppose that $F(x)$ is bounded and integrable for $a \leqq x \leqq b$ and that the variable is changed from x to u where $x = \varphi(u)$; let $u = \alpha$ when $x = a$ and $u = \beta$ when $x = b$. Both $\varphi(u)$ and $\varphi'(u)$ are to be continuous and $\varphi'(u)$ is *not to change sign* as u varies from α to β; hence $\varphi(u)$ is strictly monotonic, and as x increases from a to b, either u increases from α to β (when $\varphi'(u)$ is positive) or else u decreases from α to β (when $\varphi'(u)$ is

negative). When these conditions are satisfied u is a single-valued, monotonic function of x, say $u=\psi(x)$, and the formula for change of variable is, as before,

$$\int_a^b F(x)\,dx=\int_\alpha^\beta F[\varphi(u)]\,\varphi'(u)\,du.$$

Take the division $[\alpha, u_1, u_2, \ldots u_{n-1}, \beta]$ of the interval (α, β) and the corresponding division $[a, x_1, x_2, \ldots, x_{n-1}, b]$ of the interval (a, b) where $x_r=\varphi(u_r)$. We have

$$x_{r+1}-x_r=\varphi(u_{r+1})-\varphi(u_r)=\varphi'(v_r)(u_{r+1}-u_r) \quad \ldots\ldots\ldots(1)$$

where v_r lies between u_r and u_{r+1}; let $\xi_r=\varphi(v_r)$ so that ξ_r lies between x_r and x_{r+1}. Hence

$$\sum_{r=0}^{n-1} F(\xi_r)(x_{r+1}-x_r)=\sum_{r=0}^{n-1} F[\varphi(v_r)]\,\varphi'(v_r)(u_{r+1}-u_r). \quad \ldots\ldots(2)$$

The product of the integrable function $F[\varphi(u)]$ and the continuous function $\varphi'(u)$ is integrable and, by (1), when n tends to infinity in such a way that the length of each interval (x_r, x_{r+1}) tends to zero so does the length of each interval (u_r, u_{r+1}). If we use the definition (D_1), § (107), we now see that the formula stated above is correct.

The proof contains that for the indefinite integral; for we may suppose x and $\psi(x)$ to be put in place of b and β, where $x=\varphi(u)$ and $a \leqq x < b$.

Note. When x is defined implicitly as a function of u by an equation $f(x, u)=0$ special care is required. See *E.T.* p. 470, Ex. 6, for an illustration.

EXERCISES XIII.

1. Evaluate the integrals of e^x, $\sin(cx+c')$, $\cos(cx+c')$ over the interval (a, b) by taking a division of (a, b) into n equal parts.

2. If $0<a<b$ and $a\varrho^n=b$ so that $\varrho \to 1$ when $n\to\infty$, show that

$$\int_a^b F(x)\,dx=\mathop{\mathcal{L}}_{n\to\infty}\sum_{r=0}^{n-1} F(a\varrho^r)(a\varrho^{r+1}-a\varrho^r)$$
$$=\mathop{\mathcal{L}}_{\varrho\to 1}(\varrho-1)\sum_{r=0}^{n-1} F(a\varrho^r)a\varrho^r.$$

Deduce that

(i) $\displaystyle\int_a^b x^m\,dx=\frac{b^{m+1}-a^{m+1}}{m+1}$, $m\neq -1$; (ii) $\displaystyle\int_a^b \frac{dx}{x}=\log\frac{b}{a}$;

(iii) $\displaystyle\int_a^b \log x\,dx=(b\log b-b)-(a\log a-a)$.

3. The base BC of a triangle ABC is divided into n equal parts B_rB_{r+1}, $r=0, 1, 2, \ldots (n-1)$, (B_0, B_n denote B, C respectively), and a point P_r is taken anywhere in the segment B_rB_{r+1}; if P is any point in BC and $BP=x$, prove that, with the usual notation for the triangle ABC,

$$\underset{n\to\infty}{\mathcal{L}}\sum_{r=0}^{n-1}\frac{B_rB_{r+1}}{AP_r}=\int_0^a\frac{dx}{AP}=\log\left(\cot\frac{B}{2}\cot\frac{C}{2}\right).$$

4. If $a>0$, p a positive integer and $N=pn$, prove the following results:

(i) $$\underset{n\to\infty}{\mathcal{L}}\sum_{r=1}^{N}\frac{1}{na+r}=\log\left(1+\frac{p}{a}\right);$$

(ii) if also $b>0$,

$$\underset{n\to\infty}{\mathcal{L}}\sum_{r=1}^{N}\frac{1}{na+rb}=\frac{1}{b}\log\left(1+\frac{pb}{a}\right);$$

iii) $$\underset{n\to\infty}{\mathcal{L}}\sum_{r=1}^{N}\frac{1}{na+(2r-1)}=\tfrac{1}{2}\log\left(1+\frac{2p}{a}\right).$$

Show that the results hold if p is not integral but is greater than 1 and N such that $N\leqq np<N+1$.

5. If $S_n=\sum_{r=1}^{pn}\frac{1}{2r-1}$, $T_n=\sum_{r=1}^{qn}\frac{1}{2r}$,

where p, q are positive integers, show that

(i) $$S_n-T_n=\sum_{r=1}^{2qn}\frac{(-1)^{r-1}}{r}+\sum_{r=1}^{(p-q)n}\frac{1}{2qn+2r-1},\ p>q;$$

(ii) $$S_n-T_n=\sum_{r=1}^{2pn}\frac{(-1)^{r-1}}{r}-\sum_{r=1}^{(q-p)n}\frac{1}{2pn+2r-1},\ p<q;$$

and prove that, whether $p\gtreqless q$,

$$\underset{n\to\infty}{\mathcal{L}}(S_n-T_n)=\log 2+\frac{1}{2}\log\frac{p}{q}.$$

Deduce a theorem on the change of value of the infinite series

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots+\frac{(-1)^{n-1}}{n}+\ldots$$

produced by a certain derangement of its terms.

6. If $S_n=\sum_{r=1}^{n}\frac{2}{2n+2r-1}$, prove that $\underset{n\to\infty}{\mathcal{L}}n^2(\log 2-S_n)=\frac{1}{32}$.

7. By use of the identity

$$a^{2n}-1=(a^2-1)\prod_{r=1}^{n-1}\left(1-2a\cos\frac{r\pi}{n}+a^2\right),$$

show that

$$\int_0^\pi\log(1-2a\cos x+a^2)dx=\pi\log(a^2),\ a^2>1,$$
$$=0\qquad,\ a^2<1.$$

8. If $f(n, r)$ is a bounded, homogeneous function of n and r, of degree -1, and p a positive integer, prove that

$$\mathop{\mathcal{L}}_{n\to\infty} \sum_{r=1}^{pn} f(n, r) = \int_0^p f(1, x)\,dx.$$

9. If $u = xt$ and if $f(u)$ and all its derivatives up to and including $f^{(n)}(u)$ are continuous for $0 \leqq u \leqq x$ (or for $0 \geqq u \geqq x$ when x is negative), show that $xf'(u) = df(u)/dt$ and that

$$f(x) - f(0) = x\int_0^1 f'(u)\,dt$$
$$= \Big[-x(1-t)f'(u)\Big]_0^1 + x^2\int_0^1 (1-t)f''(u)\,dt.$$

Deduce that

$$f(x) = f(0) + \sum_{r=1}^{n-1} \frac{x^r}{r!} f^{(r)}(0) + R_n$$

where
$$R_n = \frac{x^n}{(n-1)!}\int_0^1 (1-t)^{n-1} f^{(n)}(u)\,dt,$$

and, noting that $(1-t)^{n-1} = (1-t)^{p-1}(1-t)^{n-p}$, $1 \leqq p \leqq n$, and applying the First Theorem of Mean Value, show that

$$R_n = \frac{x^n(1-\theta)^{n-p} f^{(n)}(\theta x)}{(n-1)!\,p}.$$

10. If $\varphi(x)$, $\psi(x)$ and all their derivatives up to and including $\varphi^{(n)}(x)$, $\psi^{(n)}(x)$ are continuous for $a \leqq x \leqq b$, prove that

$$\int_a^b \varphi(x)\,\psi^{(n)}(x)\,dx = \Big[F(x)\Big]_a^b + (-1)^n \int_a^b \psi(x)\varphi^{(n)}(x)\,dx$$

where
$$F(x) = \sum^{n-1} (-1)^r \varphi^{(r)}(x)\psi^{(n-r-1)}(x), \quad \varphi^{(0)}(x) = \varphi(x).$$

11. In Example 10, let $\psi(x) = (b-x)^{n-1}$ and show that

$$\varphi(b) = \sum_{r=0}^{n-1} \frac{(b-a)^r}{r!} \varphi^{(r)}(a) + R_n$$

where
$$R_n = \frac{1}{(n-1)!}\int_a^b (b-x)^{n-1}\varphi^{(n)}(x)\,dx$$
$$= \frac{(b-\xi)^{n-p}(b-a)^p\varphi^{(n)}(\xi)}{(n-1)!\,p}, \quad \begin{cases} \xi = a + \theta(b-a), \ 0 < \theta < 1, \\ b - \xi = (b-a)(1-\theta), \ 1 \leqq p \leqq n. \end{cases}$$

12. In 10 put $n+1$ for n; then prove that, if $\psi(x) = e^{\lambda x}$ and $\varphi(x)$ is a polynomial in x of degree n,

$$\int_a^b e^{\lambda x}\varphi(x)\,dx = \left[\sum_{r=0}^{n} (-1)^r \frac{e^{\lambda x}}{\lambda^{r+1}} \varphi^{(r)}(x)\right]_a^b.$$

13. Prove that

$$\frac{d}{dx}\{f^{(r-1)}(x)g^{(n-r)}(-x)\} = f^{(r)}(x)g^{(n-r)}(-x) - f^{(r-1)}(x)g^{(n-r+1)}(-x) \quad \text{...(i)}$$

where $\quad g'(-x) = \dfrac{dg(-x)}{d(-x)} = -\dfrac{dg(-x)}{dx}, \quad g^{(r)}(-x) = -\dfrac{dg^{(r-1)}(-x)}{dx}.$

In (i) give to r in succession the values $1, 2, \dots, n$ and, by adding the n values of the two members of equation (i), deduce that

$$\frac{dF(x)}{dx} = f^{(n)}(x)g(-x) - f(x)g^{(n)}(-x) \quad \text{.....................(ii)}$$

where
$$F(x) = \sum_{r=1}^{n} f^{(r-1)}(x)g^{(n-r)}(-x).$$

14. If $f(x)$ is any polynomial in x of degree less than n and $P_n(x)$ a polynomial of degree n given by

$$P_n(x) = \frac{d^n}{dx^n}\{A(x-a)^n(x-b)^n\}, \; A = \text{constant},$$

deduce from Example 13 that

$$\int_a^b f(x)P_n(x)\,dx = 0. \quad \text{.............................(i)}$$

[Let $g(x) = A(x+a)^n(x+b)^n$, so that $P_n(x) = (-1)^n g^{(n)}(-x)$; then integrate equation (ii).]

15. If $Q_n(x)$ is a polynomial of degree n such that

$$\int_a^b f(x)Q_n(x)\,dx = 0,$$

where, as before, $f(x)$ is any polynomial of degree less than n, prove that $Q_n(x) = CP_n(x)$ when C is a constant.

[We have
$$\int_a^b f(x)\{Q_n(x) - CP_n(x)\}\,dx = 0.$$

Now C may be chosen so that $Q_n - CP_n$ is of degree $n-1$ (or lower degree); let it be so chosen. Since $f(x)$ is any polynomial of degree less than n we may take $f(x) = Q_n - CP_n$ and then

$$\int_a^b \{Q_n(x) - CP_n(x)\}^2\,dx = 0.$$

But $Q_n(x) - CP_n(x)$ is continuous for $a \leqq x \leqq b$ and the integral will necessarily be positive unless $Q_n(x) - CP_n(x) = 0$ for $a \leqq x \leqq b$.

The integral in Example 14 thus expresses a characteristic property of $P_n(x)$. An important special case is the following if $m \neq n$:

$$\int_a^b P_m(x)P_n(x)\,dx = 0 \quad \text{.............................(i)}$$

Let $f(x) = P_m(x)$ if $m < n$ and $f(x) = P_n(x)$ if $m > n$.]

16. In Example 14, let $a = -1$, $b = 1$, $A = 1/2^n \,.\, n!$ so that

$$P_n(x) = \frac{1}{2^n \,.\, n!}\frac{d^n}{dx^n}\{(x^2-1)^n\}.$$

This value of $P_n(x)$ is called *The Legendre Polynomial of degree n*, or

The Legendre Coefficient of degree n; by convention, $P_0(x)$ is taken to be unity. Prove that

$$\text{(i) } P_n(1)=1,\ P_n(-1)=(-1)^n;$$

$$\text{(ii) } \int_{-1}^{1} P_m(x)P_n(x)dx=0,\ m\neq n$$
$$=2/(2n+1),\ m=n.$$

[If $m\neq n$ the result is proved in Example 15. If $m=n$ use the above form of $P_n(x)$ and note that $D_x^r.(x^2-1)^n=0$ for $r<n$ both when $x=1$ and when $x=-1$. For the general theory of these functions reference may be made to Professor MacRobert's book on *Spherical Harmonics*. As an exercise the student may show that if $f(x)$ is any polynomial of degree n,

$$f(x)=A_0x^n+A_1x^{n-1}+\ldots+A_{n-1}x+A_n,$$

it can be expressed in the form, $B_r=$constant,

$$f(x)=B_0P_n(x)+B_1P_{n-1}(x)+\ldots B_{n-1}P_1(x)+B_nP_0(x).$$

Show that B_0 may be chosen so that $f(x)-B_0P_n(x)$ is a polynomial $f_1(x)$ of degree $(n-1)$ and therefore $f(x)=B_0P_n(x)+f_1(x)$; the process may be repeated with $f_1(x)$, and so on. Further, by equation (ii), show that

$$\frac{2}{2r+1}B_{n-r}=\int_{-1}^{1} f(x)P_r(x)dx.$$

If $f(x)=xP_n(x)$, a polynomial of degree $(n+1)$, deduce from the equation

$$xP_n(x)=B_0P_{n+1}+B_1P_n+B_2P_{n-1}+\ldots+B_nP_1+B_{n+1}P_0,$$

(*a*) that $P_n(x)$ contains no power that occurs in $xP_n(x)$ so that $B_1=0$;

(*b*) by applying equation (ii) that $B_{n-r+1}=0$ if $r<n-1$; (iii) by comparing coefficients of x^{n+1} and x^{n-1} that $B_0=(n+1)/(2n+1)$, $B_2=n/(2n+1)$. Hence the relation between P_{n+1}, P_n, P_{n-1}

$$(n+1)P_{n+1}-(2n+1)xP_n+nP_{n-1}=0.]$$

17. If u and v are bounded integrable functions of x for the range $a\leqq x\leqq b$, prove that

$$\left(\int_a^b uv\,dx\right)^2\leqq\left(\int_a^b u^2dx\right)\times\left(\int_a^b v^2dx\right).$$

This inequality is known as **Schwarz's Inequality**. To prove it, let λ and μ be constants; then

$$\int_a^b (\lambda u+\mu v)^2dx=A\lambda^2+2B\lambda\mu+C\mu^2$$

where A is the integral of u^2, etc. The quadratic form cannot be negative so that $B^2\leqq AC$; the equality can only occur if u/v is constant.

18. If $F(x)$ is continuous and positive for $a\leqq x\leqq b$, prove that the product of the integral of $F(x)$ and the integral of $1/F(x)$, each taken over the interval (a, b), is least when $F(x)$ is constant.

19. If $F(x)=(1+x)^{-1}$ for $0 \leqq x \leqq 1$, and if $f(x)=F(x)$ except for the values $x=1/r$, where $r=1, 2, 3, \ldots$, prove that

$$\int_0^1 f(x)dx = \log 2.$$

20. If $F(x)$ is defined for the interval $(0, 1)$ by the condition that if r is a positive integer, $r=1, 2, 3, \ldots$

$$F(x)=(-1)^{r-1} \text{ when } (r+1)^{-1} < x < r^{-1},$$

prove that

$$\int_0^1 F(x)dx = \log 4 - 1.$$

21. If $f(x)=\int_x^{\pi/2} \log(\sin t)dt$, $0 < \eta \leqq x \leqq \frac{\pi}{2}$, prove that $f(x)$ tends to a limit when $x \to 0$.

Write $\log \sin t = \log(\sin t/t) + \log t$ and note that

$$\int \log t\,dt = t \log t - t,$$

while $t \log t \to 0$ when $t \to 0$ and $\sin t/t$ is continuous for $0 \leqq t \leqq \pi/2$.

CHAPTER X

RECTIFICATION. CURVILINEAR INTEGRALS. AREAS. REPEATED AND DOUBLE INTEGRALS. VOLUMES. SURFACES.

115. Rectification of Curves.* Let a curve AB be defined, the axes of coordinates being rectangular, by the freedom equations

$$x=f(t),\quad y=g(t),\quad t_0 \leqq t \leqq T. \quad \text{..................(1)}$$

By "the point t_r" is meant the point $A_r(x_r, y_r)$ where x_r and y_r are the values of x and y respectively for $t=t_r$; A is the point t_0 and B the point T. As t increases from t_0 to T the point (x, y) moves along the curve from A to B.

Let $[t_0, t_1, t_2, \ldots t_{n-1}, T]$, $T=t_n$, be a division of the interval (t_0, T) and let A_r be the point t_r; A is A_0 and B is A_n. Denote by A_rA_{r+1} the length of the chord A_rA_{r+1}; then, the positive value of the square root being taken throughout,

$$\begin{aligned} A_rA_{r+1} &= \sqrt{\{(x_{r+1}-x_r)^2+(y_{r+1}-y_r)^2\}} \\ &= \sqrt{\{[f(t_{r+1})-f(t_r)]^2+[g(t_{r+1})-g(t_r)]^2\}}, \end{aligned}$$

and if l_n is the sum of these chords for $r=0, 1, 2, \ldots, (n-1)$,

$$l_n=\sum_{r=0}^{n-1} A_rA_{r+1}=\sum_{r=0}^{n-1}\sqrt{\{[f(t_{r+1})-f(t_r)]^2+[g(t_{r+1})-g(t_r)]^2\}}.$$

Definition. If, when n tends to infinity in such a way that the length $(t_{r+1}-t_r)$ of each interval (t_r, t_{r+1}) tends to zero, l_n tends to a definite limit l, the curve AB is said to be *rectifiable* and the number l is defined to be the *length* of the curve AB.

If AB is a curve in three dimensions defined, with respect to rectangular axes, by the freedom equations

$$x=f(t),\quad y=g(t),\quad z=h(t),\quad t_0 \leqq t \leqq T,$$

* For a discussion of curves, areas, volumes and surfaces that involves less drastic restrictions on the defining functions the student may consult de la Vallée Poussin's *Cours d'Analyse* (2nd ed.), Vol. I, pp. 347-373 or Jordan's *Cours d'Analyse* (2nd Ed.), Vol. I, Chap. VIII.

the length of the chord A_rA_{r+1} is given by

$$A_rA_{r+1}=\sqrt{\{(x_{r+1}-x_r)^2+(y_{r+1}-y_r)^2+(z_{r+1}-z_r)^2\}}$$

and
$$l_n=\sum_{r=0}^{n-1}A_rA_{r+1}.$$

The definition just given for the length of a plane curve is taken as defining the length of a curve in space. The developments that will now be given for a plane curve are applicable with little more than verbal changes to a curve in space; they involve less complicated formulae, and the results can be at once adapted to the case of three dimensions.

It will now be proved that AB is rectifiable if $f(t)$, $f'(t)$, $g(t)$, $g'(t)$ are continuous for $t_0 \leqq t \leqq T$; these conditions are sufficient but not necessary. Jordan (see his *Cours d'Analyse*, 2nd Ed., §§ 105-108) has proved that the sufficient and necessary conditions that the curve AB should be rectifiable are that the functions $f(t)$ and $g(t)$ should be continuous and of limited variation.

By the Mean Value Theorem we have

$$f(t_{r+1})-f(t_r)=f'(\tau_r')(t_{r+1}-t_r),\quad g(t_{r+1})-g(t_r)=g'(\tau_r'')(t_{r+1}-t_r)$$

where τ_r' and τ_r'' both lie between t_r and t_{r+1}, so that

$$A_rA_{r+1}=\sqrt{\{[f'(\tau_r')]^2+[g'(\tau_r'')]^2\}}(t_{r+1}-t_r).\ \ \ldots\ldots\ldots\ldots(2)$$

This expression for A_rA_{r+1} can be put in the form

$$A_rA_{r+1}=[\sqrt{\{[f'(t_r)]^2+[g'(t_r)]^2\}}+\alpha_r](t_{r+1}-t_r)\ \ \ldots\ldots\ldots(3)$$

where α_r tends *uniformly* to zero when each difference $(t_{r+1}-t_r)$ tends to zero; the change from the form (2) to the form (3) is an essential element in the proof and the equivalence of the two forms may be shown as follows.

Let $d=\sqrt{(u^2+v^2)}$, $d_1=\sqrt{(u_1^2+v_1^2)}$ where u, u_1, v, v_1 are real numbers and d, d_1 are positive; then

$$d_1-d=\frac{d_1^2-d^2}{d_1+d}=(u_1-u)\frac{u_1+u}{d_1+d}+(v_1-v)\frac{v_1+v}{d_1+d}.$$

Now $\quad d\geqq|u|,\ d\geqq|v|,\ d_1\geqq|u_1|,\ d_1\geqq|v_1|$ so that

$$d_1+d\geqq|u_1+u|,\ d_1+d\geqq|v_1+v|,$$

and therefore

$$|d_1-d|=|\sqrt{(u_1^2+v_1^2)}-(\sqrt{u^2+v^2})|\leqq|u_1-u|+|v_1-v|.$$

In the same way we have, w and w_1 being real,

$$|\sqrt{(u_1^2+v_1^2+w_1^2)}-\sqrt{(u^2+v^2+w^2)}|\leqq|u_1-u|+|v_1-v|+|w_1-w|.$$

Now for u, v put $f'(t_r)$, $g'(t_r)$ and for u_1, v_1 put $f'(\tau_r')$, $g'(\tau_r'')$ respectively; then

$$|\sqrt{\{[f'(\tau_r')]^2+[g'(\tau_r'')]^2\}}-\sqrt{\{[f'(t_r)]^2+[g'(t_r)]^2\}}|$$
$$\leqq |f'(\tau_r')-f'(t_r)|+|g'(\tau_r'')-g'(t_r)|.$$

But $f'(t)$, $g'(t)$ are continuous and both τ_r' and τ_r'' lie between t_r and t_{r+1}; therefore, given ε, we may choose η so that

$$|f'(\tau_r')-f'(t_r)|<\tfrac{1}{2}\varepsilon,\quad |g'(\tau_r'')-g'(t_r)|<\tfrac{1}{2}\varepsilon,$$

if only $t_{r+1}-t_r<\eta$, $r=0, 1, 2, \ldots, n-1$. Hence we have

$$\sqrt{\{[f'(\tau_r')]^2+[g'(\tau_r'')]^2\}}=\sqrt{\{[f'(t_r)]^2+[g'(t_r)]^2\}}+\alpha_r,$$

where $|\alpha_r|<\varepsilon$ if $t_{r+1}-t_r<\eta$, so that α_r tends uniformly to zero when $t_{r+1}-t_r$ tends to zero.

We now have, using the form (3)

$$l_n=\sum_{r=0}^{n-1}[\sqrt{\{[f'(t_r)]^2+[g'(t_r)]^2\}}+\alpha_r](t_{r+1}-t_r)$$

and therefore, since the function $\sqrt{\{[f'(t)]^2+[g'(t)]^2\}}$ is continuous, by applying § 107, (D_3) we see that $l_n \to l$ where

$$l=\int_{t_0}^{T}\sqrt{\{[f'(t)]^2+[g'(t)]^2\}}\,dt=\int_{t_0}^{T}\sqrt{\left\{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2\right\}}\,dt. \quad \ldots(4)$$

If P is the point t on the curve and arc $AP=s$ we have, with θ as the variable of integration,

$$s=\int_{t_0}^{t}\sqrt{\left\{\left(\frac{dx}{d\theta}\right)^2+\left(\frac{dy}{d\theta}\right)^2\right\}}\,d\theta \quad \ldots\ldots(5)$$

and, for a curve in space,

$$s=\int_{t_0}^{t}\sqrt{\left\{\left(\frac{dx}{d\theta}\right)^2+\left(\frac{dy}{d\theta}\right)^2+\left(\frac{dz}{d\theta}\right)^2\right\}}\,d\theta. \quad \ldots\ldots(6)$$

Note. The class of rectifiable curves may obviously be extended to include a curve of the following type (a *composite* curve). Let the curves AC_1, C_1C_2, ..., C_mB be joined up at the points C_1, C_2, ..., C_m so as to form one curve AB and suppose that A, C_1, C_2, ..., C_m, B are the points t_0, t_1, t_2, ..., t_m, T respectively where $t_0<t_1<t_2\ldots<t_m<T$. If the functions $f(t)$, $g(t)$, $h(t)$ are continuous for $t_0 \leqq t \leqq T$ the curve AB is continuous; if also the derivatives $f'(t)$, $g'(t)$, $h'(t)$ are continuous for each of the *closed intervals* (t_0, t_1), (t_1, t_2), ..., (t_m, T) then each part AC_1, C_1C_2, ..., C_mB of the composite curve is rectifiable, and the sum of the lengths of the parts AC_1,

$C_1C_2, \dots, C_mB$ is defined to be the length of the curve AB. At the points $C_1, C_2, \dots, C_m$ one or more of the derivatives $f'(t)$, $g'(t)$, $h'(t)$ will usually be discontinuous and there will be *two tangents* at such a point, there being an angle between the backward and the forward tangents at the point (as at the point H, p. 161, Fig. 33 of the *Elem. Treat.*). The simplest example is that of a "broken line" $AC_1C_2 \dots C_mB$ in which each of the parts $AC_1, C_1C_2, \dots, C_mB$ is a straight line and no two consecutive parts are collinear.

Cor. 1. The expression for s as an integral in (6) gives

$$\frac{ds}{dt} = \sqrt{\left\{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right\}}, \quad \dots\dots\dots\dots\dots(\alpha)$$

and therefore, in terms of differentials,

$$ds^2 = dx^2 + dy^2 + dz^2 \quad \dots\dots\dots\dots\dots\dots\dots\dots(\beta)$$

(The corresponding expressions deducible from (5) are obtained by supposing dz to be identically zero.) The equation (β) holds whatever the independent variable may be.

Cor. 2. If $y = \varphi(x)$ and if $\varphi(x)$ and $\varphi'(x)$ are continuous for $a \leqq x \leqq b$, the curve has the freedom equations $x = t$, $y = \varphi(t)$, so that, replacing t by x, we have

$$ds^2 = dx^2 + dy^2 = \{1 + [\varphi'(x)]^2\}dx^2,$$

and

$$s = \int_a^x \sqrt{\{1 + [\varphi'(\xi)]^2\}}d\xi.$$

Similarly, in three dimensions, we may write

$$x = t, \quad y = \varphi(t), \quad z = \psi(t),$$

and

$$s = \int_a^x \sqrt{\{1 + [\varphi'(\xi)]^2 + [\psi'(\xi)]^2\}}d\xi.$$

Cor. 3. If for $t_0 \leqq t \leqq T$ the functions $f(t)$, $g(t)$, $h(t)$, $f'(t)$, $g'(t)$, $h'(t)$ are continuous and the derivatives are not all zero for the same value of t, the function s defined by the integral (6) is a continuous, monotonic, increasing function of t and ds/dt as well as s is continuous. Hence t is a continuous, monotonic, increasing function of s and x, y, z may therefore be taken as functions of s which, with their first derivatives, are continuous functions of s for the range $0 \leqq s \leqq l$ where l is the length of AB; that is, the curve AB may be represented by freedom equations of the form $\quad x = F(s), \quad y = G(s), \quad z = H(s).$

In this case dx/ds, dy/ds, dz/ds are the direction cosines of the tangent to the curve at the "point s," as is easily seen; the derivatives dx/dt, dy/dt, dz/dt are *proportional* to these direction cosines (for a plane curve dz may be taken to be identically zero).

Ex. 1. If the curve AB is plane and given by an equation $r=f(\theta)$ in polar coordinates, show that

$$ds^2=dr^2+r^2d\theta^2,$$

$$s=\int\sqrt{\left\{r^2+\left(\frac{dr}{d\theta}\right)^2\right\}}d\theta+\text{constant}.$$

In the equation $ds^2=dx^2+dy^2$ put $x=r\cos\theta$, $y=r\sin\theta$; then

$$dx=\cos\theta\,dr-r\sin\theta\,d\theta,\quad dy=\sin\theta\,dr+r\cos\theta\,d\theta,$$

so that

$$ds^2=dr^2+r^2d\theta^2.$$

Ex. 2. If for a curve in space the coordinates x, y, z are changed to spherical polar coordinates r, θ, φ by the transformation

$$x=r\sin\theta\cos\varphi,\quad y=r\sin\theta\sin\varphi,\quad z=r\cos\theta,$$

show, in the same way as in *Example* 1, that

$$ds^2=dr^2+r^2d\theta^2+r^2\sin^2\theta\,d\varphi^2.$$

Here $\quad dx=\sin\theta\cos\varphi\,dr+r\cos\theta\cos\varphi\,d\theta-r\sin\theta\sin\varphi\,d\varphi$

with similar expressions for dy and dz; substitute in the equation

$$ds^2=dx^2+dy^2+dz^2$$

and the result follows.

Ex. 3. If A_r and A_{r+1} are the points t_r and t_{r+1} respectively on the curve AB, prove that the ratio of the chord A_rA_{r+1} to the arc A_rA_{r+1} tends to unity when A_{r+1} tends along the arc to A_r, or, when $t_{r+1}\to t_r$.

For simplicity suppose AB a plane curve. The integral (5) gives

$$\text{arc } A_rA_{r+1}=\int_{t_r}^{t_r+1}\sqrt{\{[f'(t)]^2+[g'(t)]^2\}}\,dt=(t_{r+1}-t_r)\sqrt{\{[f'(\tau_r)]^2+[g'(\tau_r)]^2\}}$$

where τ_r lies between t_r and t_{r+1}. Also by § 115, equation (2),

$$\text{chord } A_rA_{r+1}=(t_{r+1}-t_r)\sqrt{\{[f'(\tau_r')]^2+[g'(\tau_r'')]^2\}}$$

where τ_r' and τ_r'' both lie between t_r and t_{r+1}.

But, as proved above, if $t_{r+1}-t_r<\eta$, we have

$$\left|\sqrt{\{[f'(\tau_r')]^2+[g'(\tau_r'')]^2\}}-\sqrt{\{[f'(\tau_r)]^2+[g'(\tau_r)]^2\}}\right|<\varepsilon$$

and therefore

$$\left|\frac{\text{chord } A_rA_{r+1}}{\text{arc } A_rA_{r+1}}-1\right|<\frac{\varepsilon}{\sqrt{\{[f'(\tau_r)]^2+[g'(\tau_r)]^2\}}}$$

so that, if $f'(t)$ and $g'(t)$ are not simultaneously zero for $t_r\leqq t\leqq t_{r+1}$, the ratio of the chord to the arc tends to unity when A_{r+1} tends to A_r.

This property of the ratio of chord to arc was assumed previously (*E.T.* p. 109) as an axiom; with the definition of what is meant by "the length of a curve," based on the integral, the axiom now appears as a theorem capable of proof.

Ex. 4. Prove *Example* 14, p. 361, of the *Elementary Treatise.*

Ex. 5. If a curve is represented by freedom equations $x=f(s), \dots$, the parameter s being the length of the arc from a fixed point on it up to (x, y, z), and if accents denote derivatives with respect to s ($x'=dx/ds$, $x''=d^2x/ds^2, \dots$), prove that

$$x'x''+y'y''+z'z''=0,$$

and if $\varrho^{-2}=(x'')^2+(y'')^2+(z'')^2$, show that the line whose direction cosines are $\varrho x''$, $\varrho y''$, $\varrho z''$ is perpendicular to the tangent to the curve at (x, y, z). Find also the direction cosines of the line that is perpendicular to these two lines.

Note that $(x')^2+(y')^2+(z')^2=1.$

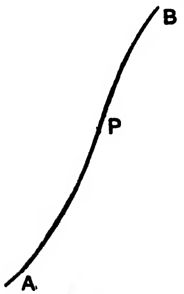

FIG. 1.

116. Curvilinear Integrals. Let $y=\varphi(x)$, where $\varphi(x)$ is single-valued and continuous for the range $a \leqq x \leqq b$, be the equation, referred to rectangular axes, of the curve APB (Fig. 1), A being the point (a, a') and B the point (b, b').

Suppose that $F(x, y)$ is a single-valued function of x and y where $y=\varphi(x)$ and form a division $[a, x_1, x_2, \dots, x_{n-1}, b]$ of the interval (a, b). Take ξ_r such that $x_r \leqq \xi_r \leqq x_{r+1}$, let $\eta_r=\varphi(\xi_r)$, and consider the sum S_n where

$$S_n=\sum_{r=0}^{n-1} F(\xi_r, \eta_r)(x_{r+1}-x_r)=\sum_{r=0}^{n-1} F\{\xi_r, \varphi(\xi_r)\}(x_{r+1}-x_r).$$

Definition. If, when n tends to infinity in such a way that the length $(x_{r+1}-x_r)$ of each sub-interval (x_r, x_{r+1}) tends to zero, S_n tends to a limit, that limit is called **an integral of $F(x, y)$ along the curve AB** (a curvilinear integral) and is denoted by the symbol

$$\int_{AB} F(x, y)\, dx.$$

The limit will certainly exist if $F(x, y)$ is a continuous function of x and y because $F\{x, \varphi(x)\}$ will be a continuous function of x for the range $a \leqq x \leqq b$, since $\varphi(x)$ is so. The sum S_n is in this case merely a particular example of the general theorem in integration, so that

$$\int_{AB} F(x, y)\, dx = \mathop{\mathcal{L}}_{n\to\infty} S_n = \int_a^b F\{x, \varphi(x)\}\, dx.$$

It is supposed in what follows that $F(x, y)$ is a continuous function of x and y.

The definition can be extended to cases in which y is not a single-valued function of x. Consider the curves [Fig. 2, (α), (β), (γ)].

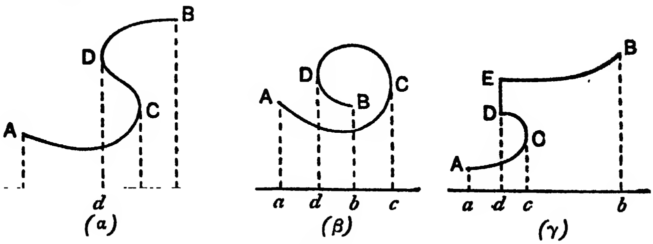

FIG. 2.

Along AC let $y=\varphi_1(x)$, $a \leqq x \leqq c$; along CD let $y=\varphi_2(x)$, $c \geqq x \geqq d$; along DB in (α) and (β) $y=\varphi_3(x)$, $d \leqq x \leqq b$; while in (γ) $x=d$ along DE and $y=\varphi_3(x)$, $d \leqq x \leqq b$ along EB.

The functions $\varphi_r(x)$ and $F\{x, \varphi_r(x)\}$, $r=1, 2, 3$, are supposed to be single-valued and continuous in the respective intervals; along DE in (γ) x is constant and the integral arising from DE is therefore zero. Thus the integral along AB is defined as the sum of the integrals along AC, CD, DB (or DE and EB), each of which has a definite value:

$$\int_{AB} F(x, y)\,dx = \int_a^c F\{x, \varphi_1(x)\}\,dx + \int_c^d F\{x, \varphi_2(x)\}\,dx + \int_d^b F\{x, \varphi_3(x)\}\,dx$$
$$= \int_{AC} F(x, y)\,dx + \int_{CD} F(x, y)\,dx + \int_{DB} F(x, y)\,dx.$$

In the same way the curvilinear integral

$$\int_{A} G(x, y)\,dy$$

is defined, x being a single-valued continuous function $\psi(y)$ say, of y when every line parallel to the x-axis meets the curve AB in only one point at most, or different single-valued continuous functions $\psi_1(y)$, $\psi_2(y)$, ... when a line parallel to the x-axis may meet AB in two or more points.

Again, the curve may be closed, like a circle or an ellipse; in this case B coincides with A and the direction of describing the curve may be indicated by taking two points C, D on the curve and using the form $ACDA$ instead of AB.

The definition may be extended to a curve in space. If the curve is defined as the intersection of the cylinders $y=\varphi(x)$,

$z=\psi(x)$ and if $F(x, y, z)$ is a single-valued function of x, y, z the curvilinear integral

$$\int_{AB} F(x, y, z)\,dx$$

means
$$\int_a^b F\{x, \varphi(x), \psi(x)\}\,dx$$

where a, b are the x-coordinates of A, B respectively. Corresponding definitions hold for integrals with respect to y and z.

117. Area. Let $y=F(x)$, where $F(x)$ is single-valued and continuous for the range $a \leqq x \leqq b$, be the equation, referred to rectangular axes, of a curve CD and let AC, BD be the ordinates at C, D so that $AC=F(a)$ and $BD=F(b)$. The area of a polygon—that is, a closed plane figure bounded by straight lines—has a definite measure, but when the boundary of a closed figure consists in whole or in part of curved lines the method by which the measure of a polygon is determined is no longer applicable and the measure of the area of such a figure needs definition. The measure may be defined in the following way.

First, suppose that $F(x)$ is positive for $a \leqq x \leqq b$ and take a division $[a, x_1, x_2, \dots, x_{n-1}, b]$ of the interval (a, b). Let M_r and m_r be the maximum and minimum values of $F(x)$ in the sub-interval (x_r, x_{r+1}) of length $h_r(=x_{r+1}-x_r)$, and let A_rP_r and $A_{r+1}P_{r+1}$ be the ordinates $F(x_r)$ and $F(x_{r+1})$. The figure $A_rA_{r+1}P_{r+1}P_r$—where P_rP_{r+1} is the arc of the curve CD between P_r and P_{r+1}—lies between the two rectangles whose areas are M_rh_r and m_rh_r respectively. Thus the figure $ABDC$ lies between two sets of rectangles whose total areas are S and s respectively where

$$S=\sum_{r=0}^{n-1} M_rh_r \text{ and } s=\sum_{r=0}^{n-1} m_rh_r.$$

When $n\to\infty$ and each $h_r\to 0$ the numbers S and s have a limit which is the same for each, namely, the integral

$$\int_a^b F(x)\,dx.$$

This integral is defined to be the *measure of the area ABDC*, or more simply, *the area ABDC*, when it is obvious that its measure is in question.

The extension to other cases then follows exactly as is shown in §§ 80 and 128 of the *Elementary Treatise*; the rule for determining the sign of the area, § 80, p. 187, and § 128, p. 318, should be noted.

118. Area of a Closed Curve. Let $ACDA$ be a closed curve without a double point, and let its freedom equations be

$$x=f(t),\quad y=g(t), \quad\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

the point (x, y) describing the curve as t varies from t_0 to T. If t_1 and t_2 are unequal values of t and if both lie between t_0 and T the points t_1 and t_2 will be different because the curve has no double point; on the other hand, the points t_0 and T are the same.

If it be assumed further that $f'(t)$ and $g'(t)$ are continuous in the interval (t_0, T) it may be proved, as in § 128 of the *Elementary Treatise*, that when the point (x, y) moves round the curve in the positive direction the area enclosed by the curve is given by each of the three integrals

$$\int_{t_0}^{T} x\frac{dy}{dt}\,dt,\ -\int_{t_0}^{T} y\frac{dx}{dt}\,dt,\ \tfrac{1}{2}\int_{t_0}^{T}\left(x\frac{dy}{dt}-y\frac{dx}{dt}\right)dt \quad\ldots\ldots\ldots(2)$$

The functions $f(t)$, $f'(t)$ and $g(t)$, $g'(t)$ are by hypothesis continuous and therefore the curve $ACDA$ is rectifiable.

Next suppose that the closed curve is a *composite curve*, that is, as explained in the *Note*, § 115, a curve formed by joining up the curves AC_1, C_1C_2, $C_2C_3, \ldots$, $C_{m-1}C_m$, C_mA at the points $C_1, C_2, \ldots, C_m, A$. If A is the point t_0 (or T since the curve is closed) and $C_1, C_2, \ldots, C_m$ the points $t_1, t_2, \ldots, t_m$ where $t_0<t_1<t_2< \ldots <t_m<T$, and if the functions $f(t)$ and $g(t)$ give the coordinates x and y of any point on the curve, we assume (i) that $f(t)$ and $g(t)$ are continuous in the closed interval (t_0, T) and (ii) that $f'(t)$ and $g'(t)$ are continuous in each of the closed intervals (t_0, t_1), $(t_1, t_2), \ldots, (t_m, T)$. The curve is rectifiable and the area enclosed by the curve will still be given by the integral (2).

One or more of the curves AC_1, $C_1C_2, \ldots$ may be straight lines; in particular they may be segments parallel to one or other of the coordinate axes. For example, AC_1 might be part of the x-axis, C_1C_2 and AC_m parallel to the y-axis while the abscissae of the points $C_3, C_4, \ldots, C_{m-1}$ might all lie between

the abscissae of A and C_1; in this case the closed area is "the area under the curve $C_m C_2$," just as (*E.T.* p. 185) the area $ABDC$ is the area under the curve CD.

It is sometimes useful to define $f(t)$ and $g(t)$ for values of t that lie outside the interval (t_0, T); the definition is simply to make them periodic, with period $(T - t_0)$, so that

$$f[n(T-t_0)+t] = f(t), \quad g[n(T-t_0)+t] = g(t),$$

where n is any positive or negative integer.

Again, if the origin of coordinates be changed and the axes turned through an angle α the old and new coordinates, (x, y) and (ξ, η) are connected by equations of the form

$$x = \xi \cos \alpha - \eta \sin \alpha + a, \quad y = \xi \sin \alpha + \eta \cos \alpha + b.$$

Now since the (closed) curve is rectifiable the length s of the arc, from a fixed point on the curve up to the variable point (x, y), may be taken instead of t in the freedom equations; s will be taken to be positive when measured in the direction that is taken as the positive direction of motion of the point (x, y). If l is the length of the curve, its area is given by the integral

$$\tfrac{1}{2}\int_0^l \left(x\frac{dy}{ds} - y\frac{dx}{ds}\right)ds,$$

and it is easily proved that this is equal to

$$\tfrac{1}{2}\int_0^l \left(\xi\frac{d\eta}{ds} - \eta\frac{d\xi}{ds}\right)ds,$$

so that the number that measures the area is independent (as it should be) of any particular coordinate axes.

Note. **Conditions to be satisfied by a curve.** It will be assumed in all that follows that a curve may be defined, in the manner illustrated, by freedom equations $x = f(t)$, $y = g(t)$ where $f(t)$ and $g(t)$ are continuous and the derivatives $f'(t)$ and $g'(t)$ *in general* continuous—that is, continuous except for a finite number of values of t. A curve as thus defined is both rectifiable and quadrable—that is, any arc of the curve has a definite length and the area enclosed by the curve (if it be closed) or the area bounded by the x-axis, an arc AC (for which the ordinate is single-valued) and the ordinates at A and C is measured by a definite number.

EXERCISES XIV.

1. A curve is given by the freedom equations

$$x = a\cos t - \tfrac{1}{3}(a-b)\cos^3 t,\quad y = b\sin t + \tfrac{1}{3}(a-b)\sin^3 t;$$

the length of the curve, measured from the point $t=0$, is

$$\tfrac{1}{2}(a+b)t - \tfrac{1}{2}(a-b)\cos t\sin t.$$

2. The length of that arc of the curve

$$4(x^2+y^2) = 3a^{\frac{4}{3}}x^{\frac{2}{3}} + a^2$$

which lies in the first quadrant is $\frac{3}{2}a$ and the length of the whole curve is $6a$.

3. The length of each of the following curves is $x+z$, when the point on each curve from which the length is measured is properly chosen:

(i) $2ay = x^2,\ 6a^2z = x^3$;

(ii) $y = \sqrt{(a^2-x^2)},\ z = \frac{1}{4}a\log\dfrac{a+x}{a-x} - \frac{1}{2}x$;

(iii) $y = a\sin^{-1}\left(\dfrac{x}{a}\right),\ z = \frac{1}{4}a\log\dfrac{a+x}{a-x}$. (Schlömilch)

4. The whole length of the curve given by the freedom equations

$$x = \frac{(t+1)^2}{3t^2+1}a,\quad y = \frac{(t-1)^2}{3t^2+1}a,\quad z = \frac{t^2-1}{3t^2+1}a$$

is $\dfrac{2\sqrt{6}}{3}\pi a$.

5. If $x = a\cos\theta,\ y = a\sin\theta,\ z = c\theta$, show that $s = \sqrt{(a^2+c^2)}\,.\,\theta$.

6. If $x = at\cos t,\ y = at\sin t,\ z = ct$, show that

$$s = \tfrac{1}{2}t(a^2+c^2+a^2t^2)^{\frac{1}{2}} + \frac{a^2+c^2}{2a}\log\left\{\frac{at+(a^2+c^2+a^2t^2)^{\frac{1}{2}}}{(a^2+c^2)^{\frac{1}{2}}}\right\}.$$

7. If $x = a\cosh t\cos t,\ y = a\cosh t\sin t,\ z = at$ then $s = \sqrt{2}\,.\,a\sinh t$.

8. Given that

$$x = a\cos\frac{a\theta}{b}\cos\theta + b\sin\frac{a\theta}{b}\sin\theta,\quad z = \sqrt{(b^2-a^2)}\left(1-\cos\frac{a\theta}{b}\right),$$

$$y = a\cos\frac{a\theta}{b}\sin\theta - b\sin\frac{a\theta}{b}\cos\theta,$$

show that $s = bz/a$. Prove that a tangent to the curve makes a constant angle with the z-axis.

9. On the sphere given by the freedom equations

$$x = a\sin\theta\cos\varphi,\quad y = a\sin\theta\sin\varphi,\quad z = a\cos\theta,$$

a curve is determined by the equation $\sin\theta\cosh n\varphi = 1$; show that, if $n = \cot\alpha$, the length of an arc, measured from the point $(0, 0, a)$ is $a\theta\sec\alpha$ and that the curve cuts the curves $\varphi=$const. at a constant angle (α).

10. The area enclosed by the curve
$$a^2(x^2+y^2)^2(b^2x^2+a^2y^2)=(a^2-b^2)^2b^2x^4$$
is $\pi b(a-b)^2(2a+b)/2a^2$.

11. The area enclosed by the curve
$$a^2(b^2x^2+a^2y^2)^3=(a^2-b^2)^2b^6x^4$$
is $3\pi b(a^2-b^2)^2/8a^3$.

12. Show that the curve given by the equation
$$x^4+4x^2y^2-6a^2x^2+a^4=0$$
consists of two ovals and that the area of each oval is πa^2.

13. A curve is given by the equations
$$x=a\cos\theta,\ y=a(2+\sin\theta)\sin^2\theta/(3+\cos^2\theta);$$
show that the area enclosed by it is $(16-9\surd 3)\pi a^2/\surd 3$.

14. The area enclosed by the curve $(x^2+a^2)^5y^2=a^{10}(a^2-x^2)$ can be expressed in terms of Gamma Functions.

15. $\int_{ABA'}(x^2+y^2)dx=-\frac{4}{3}a(a^2+2b^2)$ where ABA' is the upper half of the ellipse $x^2/a^2+y^2/b^2=1$, A being the end of the major axis $A'A$ that lies on the positive side of the origin.

16. $\int xy\,dx$ round the cardioid $r=a(1-\cos\theta)$ is $5\pi a^3/4$.

17. $\int yz\,dx$ along the curve defined by the equations
$$x=a\cos\theta,\ y=a\sin\theta,\ z=c\theta$$
from the point $(a, 0, 0)$ to the point $(a, 0, 2\pi c)$ is $-\pi^2a^2c$.

18. $\int(y\,dx+z\,dy+x\,dz)$ along the curve in which the plane $x+z=R$ intersects the sphere $x^2+y^2+z^2=R^2$ is equal to $-\pi R^2/\surd 2$; the path begins at $(R, 0, 0)$ and lies at first in the positive octant of the sphere.

19. $\int[(y^2+z^2)dx+(z^2+x^2)dy+(x^2+y^2)dz]=-2\pi ab^2$, when the path of integration is that part for which $z\geqq 0$ of the intersection of the surfaces
$$x^2+y^2+z^2=2ax \text{ and } x^2+y^2=2bx,\ a>b>0;$$
the path begins at the origin and runs at first in the positive octant.

119. Integral as Function of a Parameter. When the integrand contains numbers $y, z, \ldots$ besides the variable of integration x the integral will usually be a function of these numbers or *parameters* as they are often called when they become subsidiary variables; these parameters are constants so far as integration with respect to x is concerned, but they

usually vary within some prescribed range. Some properties of the integral as a function of one parameter, y say, will now be considered; the discussion when there are more parameters than one may be carried out on similar lines.

The integrand $F(x, y)$ is assumed to be single-valued and *bounded* and the integral of $F(x, y)$ over (a, b) will be denoted by $f(y)$ so that

$$f(y)=\int_a^b F(x, y)\,dx.$$

Further, the function $F(x, y)$ is supposed to be defined for a region bounded by a closed curve, the boundary being included in the region, and every curve is assumed to be rectifiable (§ 115, Note).

The property of $f(y)$ that will be first considered is its continuity and the following simple examples give some suggestions regarding the behaviour of $f(y)$ when discontinuities occur in $F(x, y)$.

Ex. 1. Let $F(x, y)$ be defined for the square bounded by the lines $x=0$, $x=a$ and $y=0$, $y=a$ as follows: $F(x, y)=x^2+y^2$ for all points of the square except for the sides $x=0$, $x=a$ and the diagonal $x=y$, in which cases $F(x, y)=0$.

By integration it is found at once that $f(y)=ay^2+\frac{1}{3}a^3$, so that though $F(x, y)$ is discontinuous for all points (except the origin) on the lines $x=0$, $x=a$ and $x=y$ the integral $f(y)$ is a continuous function of y in the closed interval $(0, a)$.

The three lines $x=0$, $x=a$ and $x=y$ are called *lines of discontinuity* for the function $F(x, y)$.

Ex. 2. The same as *Ex.* 1 except that $F(x, y)=0$ when $y=a$ so that the line $y=a$ is a fourth line of discontinuity.

In this case $f(y)=ay^2+\frac{1}{3}a^3$ when $0\leqq y<a$ but $f(y)=0$ when $y=a$ so that $f(y)$ is discontinuous at the end a of the interval $(0, a)$.

If $F(x, y)=0$ when $y=b<a$, as well as on the other four lines of discontinuity $f(y)$ would be discontinuous at b as well as at a.

Ex. 3. For the cube bounded by the planes $x=0$, $x=a$, $y=0$, $y=a$, $z=0$, $z=a$, let $F(x, y, z)=x^2+y^2+z^2$ except for points in the planes $x=y$ and $x=z$ in which cases $F(x, y, z)=0$. If $f(y, z)$ is the integral of $F(x, y, z)$ with respect to x over $(0, a)$, show that $f(y, z)$ is a continuous function of y and z in the square given by $y=0$, $y=a$ and $z=0$, $z=a$.

Consider the continuity of $f(y, z)$ when the planes $y=a$ and $z=a$ are also planes of discontinuity for $F(x, y, z)$.

Discontinuities. The integrand $F(x, y)$ will be assumed to be in general continuous in its region of definition but it may

be discontinuous at a finite or at an infinite number of points in the region. When the number of points of discontinuity is infinite they will be restricted by the condition that they will be assumed to lie on a finite number of curves (including straight lines) none of which is parallel to the x-axis or can be cut by a straight line parallel to the x-axis in more than a finite number of points. At a point (x_1, y_1) of discontinuity $F(x, y)$ will be assumed to be defined—that is, $F(x_1, y_1)$ will have definite value; the precise value does not matter so long as it is finite.

When the discontinuities satisfy the above conditions they may be said to be *normal*; if a line of discontinuity is parallel to the x-axis this case must be explicitly stated and discussed.

Of course if x is the parameter and y the variable of integration the normal discontinuities of $F(x, y)$ would exclude lines of discontinuity parallel to the y-axis while no line of discontinuity would be met by a parallel to the y-axis in more than a finite number of points.

Notation. When the region of definition is the rectangle R bounded by the lines $x=a$, $x=b$ and $y=a'$, $y=b'$, the region will, for brevity, be sometimes called "the rectangle $R(a, a'\,;\, b, b')$; the points (a, a') and (b, b') are opposite vertices of the rectangle.

120. Continuity with respect to a Parameter. Suppose first that $F(x, y)$ is defined for the rectangle $R(a, a'\,;\, b, b')$.

THEOREM I. *Let $F(x, y)$ be integrable with respect to x for every fixed value of y in R. If $F(x, y)$ is continuous in R or has only normal discontinuities in R then $f(y)$ where*

$$f(y)=\int_a^b F(x, y)\,dx$$

is a continuous function of y for the range $a' \leq y \leq b'$.

Case (i), $F(x, y)$ continuous. Let c and $c+k$ be two values of y in R; then

$$f(c+k)-f(c)=\int_a^b \{F(x, c+k)-F(x, c)\}\,dx.$$

Now $F(x, y)$ is continuous and therefore, by the property of uniform continuity, there is a positive number η such that,

whatever value x may take in (a, b), ε having the usual meaning,

$$|F(x, c+k) - F(x, c)| < \varepsilon \text{ if } |k| < \eta,$$

and therefore $|f(c+k) - f(c)| < \varepsilon(b-a)$ if $|k| < \eta$. Hence $f(y)$ is continuous at c where c is any number in (a', b') so that $f(y)$ is continuous in the (closed) interval (a', b').

Case (ii), $F(x, y)$ discontinuous.

Let there be one curve of discontinuity DD' (Fig. 3) and let it be met by a parallel to the x-axis in only one point at most; say that $y=c$ meets DD' where $x=\alpha$.

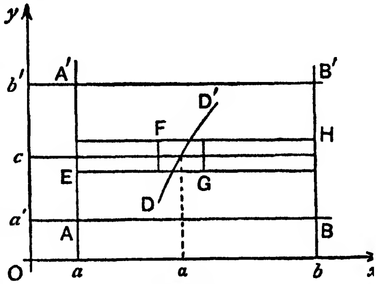

FIG. 3.

If δ is an arbitrarily small positive number it is possible to choose η so that if (x, y) is in either of the rectangles $EF(a, c-\eta\,;\ \alpha-\delta, c+\eta)$ and $GH(\alpha+\delta, c-\eta\,;\ b, c+\eta)$ the function $F(x,y)$ is continuous. If $|F(x, y)| < M$ in the small rectangle FG whose centre is the point (α, c) the contribution from that region to the difference $|f(c+k) - f(c)|$ is less than $2M \times 2\delta$ when $|k| < \eta$. Since δ, and therefore $4M\delta$, is arbitrarily small it now follows by Case (i) (because $F(x, y)$ is continuous in the rectangles EF and GH) that $f(y)$ is continuous at c.

If there were more lines of discontinuity than one the line $y=c$ would meet these in a finite number of points at most, say the points which had $\alpha_1, \alpha_2, \ldots \alpha_m$ respectively for abscissae. The neighbourhood of each of these points (α_r, c) could be treated as has been done in the case of the point (α, c); the contribution to $|f(c+k) - f(c)|$ from these neighbourhoods would be less than $(m \times 4M\delta)$ when $|k| < \eta$, and therefore would be arbitrarily small. Outside the small rectangles with centres (α_r, c) the function $F(x, y)$ is continuous so that $f(y)$ is continuous.

Note. It is now clear that there will be no loss of generality in assuming that there is only one line of discontinuity and, as a rule, the proof will be given for only one line.

Suppose next that $F(x, y)$ is defined for all points inside or on the boundary of an area D, bounded by a closed curve C.

THEOREM II. *Let $F(x, y)$ be integrable with respect to x for every fixed value of y in D. If $F(x, y)$ is continuous in D or has only normal discontinuities in D then $f(y)$ is a continuous function of y.*

The curve C will be assumed to be such that it can be cut by a line parallel to either axis in not more than two points; if this condition is not satisfied it will be assumed that the area may be divided into a finite number of parts for each of which the condition is satisfied, so that when the theorem has been proved for one part it will hold for the region composed of the sum of the parts. See, for example, Fig. 11 (a); the lines PQ, RS and TU divide the area into three parts each of which satisfies the required condition.

Let C be the curve $EFGH$ (Fig. 4). The curve lies wholly between the lines $x=a$, $x=b$ and $y=a'$, $y=b'$ and we suppose that the equation of EHG is $x=\varphi_1(y)$ and that of EFG is $x=\varphi_2(y)$, so that $\varphi_1(y)$ and $\varphi_2(y)$ are each single-valued, continuous functions of y for the range $a' \leqq y \leqq b'$.

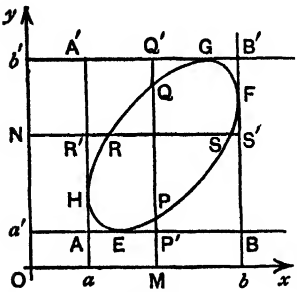

FIG. 4.

The theorem to be proved can be reduced to the Theorem I in the following way. Let the function $F_1(x, y)$ be defined so that $F_1(x,y)=F(x, y)$ for all points inside or on the boundary of the area $EFGH$, but $F_1(x, y)=0$ for all other points in the rectangle $ABB'A'$. If the curve C, that is, $EFGH$ is taken as a line of discontinuity for $F_1(x, y)$ the discontinuities of $F_1(x, y)$ are the same as those of $F(x, y)$ and in addition those that lie on C. The function $f(y)$ where

$$f(y)=\int_a^b F_1(x, y)dx$$

is continuous for $a' \leqq y \leqq b'$ by Theorem I. But if $y=ON=c$,

$$f(c)=\int_{NR'}^{NS'} F_1(x, c)dx=\int_{NR}^{NS} F(x, c)dx=\int_{\phi_1(c)}^{\phi_2(c)} F(x, c)dx$$

because $F_1(x,c)=0$ if $NR' \leqq x < NR$ or if $NS < x \leqq NS'$ and $F_1(x,c)=F(x,c)$ if $NR \leqq x \leqq NS$. Hence $f(y)$ is continuous at c and c is any number in (a', b').

The curve C may of course consist in part of straight lines. For example, C might be formed by the arc FGH and the straight lines HA, AB, BF.

Ex. 1. If $F(x, y, z)$ is bounded when (x, y, z) is any point of the cube given by $x=0$, $x=a$, $y=0$, $y=a$, $z=0$, $z=a$, and if

$$f(y, z)=\int_0^a F(x, y, z)dx,$$

show that $f(y, z)$ is a continuous function of y and z if $F(x, y, z)$ is continuous or if, when discontinuous, its discontinuities all lie in the planes $x=y$ and $x=z$.

Ex. 2. If $F(x, y, z)$ is bounded when (x, y, z) is any point in the tetrahedron whose vertices are the points $(0, 0, 0)$, (a, a, a), $(a, a, 0)$ and $(0, a, 0)$ and if

$$f(y, z)=\int^y F(x, y, z)dx,$$

show that $f(y, z)$ is continuous if $F(x, y, z)$ is continuous or if, when discontinuous, its discontinuities lie in the plane $x+y+z=a$.

121. Differentiation and Integration. Consider first the differentiation of $f(y)$.

Differentiation. With the notation of the preceding article let $F(x, y)$ and the partial derivative $\partial F/\partial y$ be continuous functions of x and y in the rectangle R; then $f(y)$ has a derivative given by the equation

$$\frac{df(y)}{dy}=\int_a^b \frac{\partial F(x, y)}{\partial y}dx, \quad\text{.......................}(1)$$

that is, given "by differentiating with respect to y under the sign of integration."

If y and $y+k$ are both in (a', b') we have

$$\frac{f(y+k)-f(y)}{k}=\int_a^b \frac{F(x, y+k)-F(x, y)}{k}dx=\int_a^b \frac{\partial F(x, y_1)}{\partial y}dx,$$

where, by the Mean Value Theorem (§ 34), y_1 lies between y and $y+k$. But by hypothesis $\partial F/\partial y$ is a continuous function of x and y, and therefore η can be chosen so that, for $a \leqq x \leqq b$ and $a' \leqq y \leqq b'$.

$$\left|\frac{\partial F(x, y_1)}{\partial y}-\frac{\partial F(x, y)}{\partial y}\right|<\varepsilon \text{ if } |k|<\eta.$$

If we now write

$$\frac{f(y+k)-f(y)}{k}=\int_a^b \frac{\partial F(x, y)}{\partial y}dx+\int_a^b\left\{\frac{\partial F(x, y_1)}{\partial y}-\frac{\partial F(x, y)}{\partial y}\right\}dx$$

equation (1) follows at once.

If a and b are not constants but differentiable, and therefore *continuous*, functions of y, write $f(y)$ in the form $f(y, a, b)$ and then the total derivative is given by

$$\frac{df}{dy}=\frac{\partial f}{\partial y}+\frac{\partial f}{\partial a}\frac{da}{dy}+\frac{\partial f}{\partial b}\frac{db}{dy}.$$

But
$$\frac{\partial f}{\partial a}=-F(a, y),\ \frac{\partial f}{\partial b}=F(b, y),$$

and $\partial f/\partial y$ is given by equation (1) so that

$$\frac{df}{dy}=\int_a^b\frac{\partial F(x, y)}{\partial y}dx-F(a, y)\frac{da}{dy}+F(b, y)\frac{db}{dy}. \quad \text{.........(2)}$$

Cor. A curvilinear integral is reducible to an ordinary integral and therefore the above investigation applies to the integral of $F(x, y, \lambda)$ when F and $\partial F/\partial\lambda$ are continuous functions of x, y, λ. If the points A and B are fixed

$$\frac{d}{d\lambda}\int_{AB}F(x, y, \lambda)dx=\int_{AB}\frac{\partial F(x, y, \lambda)}{\partial\lambda}dx. \quad \text{............(3)}$$

Integration. The function $f(y)$ may, as we have seen, be continuous for $a'\leqq y\leqq b'$ even though $F(x, y)$ is discontinuous in R, but for the present $F(x, y)$ will be supposed to be continuous in R; $f(y)$ is therefore integrable over (a', b') and the integral may be written

$$\int_{a'}^{b'}f(y)dy=\int_{a'}^{b'}\left\{\int_a^b F(x, y)dx\right\}dy=\int_{a'}^{b'}dy\int_a^b F(x, y)dx \quad \text{......(4)}$$

the latter form being the usual one. The two-fold integration gives a "repeated" (or "iterated") integral, with the meaning that "$F(x, y)$ is to be first integrated as to x, the parameter or variable y being treated as a constant in this first integration, and then the result of the x-integration is to be integrated with respect to y."

It will now be proved that if $F(x, y)$ is continuous in R we may interchange the order of integration and write

$$\int_{a'}^{b'}f(y)dy=\int_{a'}^{b'}dy\int_a^b F(x, y)dx=\int_a^b dx\int_{a'}^{b'}F(x, y)dy \quad \text{......(5)}$$

so that the integral of $f(y)$ is found "by integrating under the sign of integration."

The integrals of $F(x, y)$ over (a, b) with respect to x and over (a', b') with respect to y exist and are continuous functions of y and x respectively so that both of the repeated integrals

in (5) exist. Let t be any chosen number in (a, b) and put t in place of b in both of the repeated integrals in (5); both of these integrals will be zero if $t=a$, and therefore they will be equal if their derivatives with respect to t are equal.

Let $\varphi(t, y)$ and $\psi(x)$ be defined as follows:

$$\int_a^t F(x, y)dx=\varphi(t, y), \quad \int_{a'}^{b'} F(x, y)dy=\psi(x),$$

then we have

$$\frac{d}{dt}\int_{a'}^{b'} dy \int_a^t F(x, y)dx = \frac{d}{dt}\int_{a'}^{b'} \varphi(t, y)dy = \int_{a'}^{b'} \frac{\partial\varphi(t, y)}{\partial t}\,dy$$

and therefore
$$= \int_{a'}^{b'} F(t, y)dy.$$

Again

$$\frac{d}{dt}\int_a^t dx\int_{a'}^{b'} F(x, y)dy = \frac{d}{dt}\int_a^t \psi(x)dx=\psi(t)$$

and therefore
$$= \int_{a'}^{b'} F(t, y)dy.$$

Thus
$$\int_{a'}^{b'} dy\int_a^t F(x, y)dx = \int_a^t dx\int_{a'}^{b'} F(x, y)dy,$$

and t is any number in (a, b), so that b may be put for t.

It must be noted that this change in the order of integration without change in the value of the repeated integral assumes that the limits a, b and a', b' *are constants*.

For an extension of the conditions on which this change of order of integration is allowable see § 126.

Ex If $I=\int_0^{\frac{\pi}{2}} \frac{d\theta}{a\cos^2\theta+b\sin^2\theta}=\frac{\pi}{2\sqrt{(ab)}}$, $a>0$, $b>0$,

show by differentiating I with respect to a and b that

(i) $\displaystyle\int_0^{\frac{\pi}{2}} \frac{\cos^2\theta\, d\theta}{(a\cos^2\theta+b\sin^2\theta)^2}=\frac{\pi}{4}\cdot\frac{1}{a\sqrt{(ab)}}$;

(ii) $\displaystyle\int_0^{\frac{\pi}{2}} \frac{\cos^2\theta\sin^2\theta\, d\theta}{(a\cos^2\theta+b\sin^2\theta)^3}=\frac{\pi}{16}\,\frac{1}{ab\sqrt{(ab)}}$

Here (i) $\displaystyle\frac{\partial I}{\partial a}=\int_0^{\frac{\pi}{2}} \frac{-\cos^2\theta\, d\theta}{(a\cos^2\theta+b\sin^2\theta)^2}=-\frac{\pi}{4}\,\frac{1}{a\sqrt{(ab)}}$;

(ii) $\displaystyle\frac{\partial^2 I}{\partial b\,\partial a}=\int_0^{\frac{\pi}{2}} \frac{2\cos^2\theta\sin^2\theta\, d\theta}{(a\cos^2\theta+b\sin^2\theta)^3}=\frac{\pi}{8}\,\frac{1}{ab\sqrt{(ab)}}$, etc.

The usual forms are obtained by putting a^2 and b^2 for a and b respectively after differentiation.

122. Double Integrals. Let A be an area bounded by a closed curve C and $F(x, y)$ a function of two independent variables x and y that is single-valued and bounded in A; the integral of $F(x, y)$ over the area A will now be defined and, as the preliminary considerations that lead to the definition are in substance identical with those on which the definition of the integral of a function of a single variable is based, the statement of them may be made in a condensed form.

Let a division, D say, of the area A be made by dividing it into n elementary areas $\sigma_1, \sigma_2, \ldots, \sigma_n$, which may for brevity be called *meshes*; for example, the meshes may be formed by drawing two sets of curves that cover the area like a net (Fig. 5). The longest chord d_r of the mesh σ_r—that is, the upper limit of the distance between two points on the boundary of σ_r—will be called the *diagonal* of the mesh, and the area σ_r will tend to zero in all its dimensions when d_r tends to zero. Obviously $\sigma_r < d_r^2$; there can be no ambiguity in using the symbols σ_r, A to denote both the areas and their measures

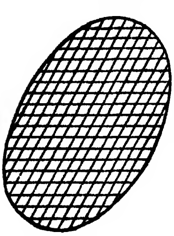

FIG. 5

Now let M, m and M_r, m_r be the upper and lower bounds of $F(x, y)$ in A and in σ_r respectively; the sums S and s where

$$S = \sum_{r=1}^{n} M_r\sigma_r, \quad s = \sum_{r=1}^{n} m_r\sigma_r$$

are called the upper and lower sums respectively for the function $F(x, y)$ and the division D of the area A.

The properties 1 ... 5 stated for the sums S and s in § 102 are also true in this case, the nomenclature used in the discussion being suitably interpreted. Thus, the division D_1 of the area A is *consecutive* to the division D if it is formed from D by dividing one or more of its meshes into two or more smaller meshes. The division D_2 is formed by *superposition* of the divisions D and D_1 when the net for the division D_2 contains all the lines that are present in the nets for the divisions D and D_1; of course, when a mesh of D_1 coincides completely with a mesh of D that mesh appears only once in D_2.

The change from the work of § 102 to the present case is simply made by substituting "area A" and "mesh σ_r" for

"interval (a, b)" and "sub-interval (x_r, x_{r+1})"; for "$h_r < h$" such a phrase as "$\sigma_r < d^2$" or "$d_r < d$" will be used. As an example, consider the property 3, § 102.

Let σ_r be a mesh of the net for the division D and S the upper sum for that division. If σ_r is divided into two or more meshes $\sigma_r', \sigma_r'', \ldots$ in which the upper bounds of $F(x, y)$ are $M_r', M_r'', \ldots$ respectively, and if S_1' is the upper sum for the new division D_1' which is consecutive to D (σ_r alone being divided), then

$$S - S_1' = M_r\sigma_r - (M_r'\sigma_r' + M_r''\sigma_r'' + \ldots)$$
$$= (M_r - M_r')\sigma_r' + (M_r - M_r'')\sigma_r'' + \ldots$$

since $\sigma_r = \sigma_r' + \sigma_r'' + \ldots$. But $M_r', M_r'', \ldots$ are each less than or, at most, equal to M_r and $M_r \leqq M$ while $M_r', M_r'', \ldots$ are each not less than m; therefore

$$0 \leqq S - S_1' \leqq (M - m)\sigma_r < (M - m)d^2 \text{ if } \sigma_r < d^2 \quad \ldots\ldots\ldots(\alpha)$$

If μ of the meshes ($\mu \leqq n$) are each divided into two or more meshes, thus forming a division D_1 consecutive to D, and if S_1 is the new value of S, then

$$0 \leqq S - S_1 < \mu(M - m)d^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(\beta)$$

when the diagonal of each mesh in the division D is less than d.

It is therefore merely a repetition of § 103 to show that S and s tend respectively to the lower limit L and the upper limit l when n tends to infinity in such a way that the diagonal of each mesh tends to zero. On account of its importance Darboux's Theorem will be stated explicitly.

Darboux's Theorem. *If D is a division of the area A for which the upper and lower sums are S and s respectively then, to any given ε, where ε is an arbitrarily small positive number, there corresponds a positive number d such that*

$$0 \leqq S - L < \varepsilon, \quad 0 \leqq l - s < \varepsilon,$$

when the diagonal of each mesh is less than d.

Or, $S \to L$ and $s \to l$ when n tends to infinity in such a way that the diagonal of each mesh tends to zero.

Note. For the special but important case in which $L = l$ the theorem may be put in the form: If *one* division of the area A can be found for which $S - s < \varepsilon$ then S and s tend to limits which are the same for both.

For, (i) S and s are bounded and $S \geqq s$; (ii) S is monotonic

and decreasing and therefore tends to a limit L while s is monotonic and increasing and therefore tends to a limit l. The condition that $L=l$ is then simply $S-s<\varepsilon$.

Double Integral. *Definitions.* The limits L and l are called respectively the upper and the lower double integrals of $F(x, y)$ over the area (or field, or region) A and are denoted by the symbols

$$L=\overline{\int_A} F(x, y)d\sigma, \quad l=\underline{\int_A} F(x, y)d\sigma.$$

If $L=l$, the common limit of S and s is called the double integral of $F(x, y)$ over the area (or field, or region) A, and is denoted by the symbol

$$\int_A F(x, y)d\sigma.$$

The symbol $d\sigma$ corresponds to the elementary area σ_r, and is often called "the element of area"; the letter A annexed to the symbol of integration indicates the area over which the integration is taken. Other notations will be given later.

123. Division of the Area. In the division D of the area A the meshes σ_r may be of any shape; the limits of S and s exist provided the diagonal of each mesh tends to zero. The division of the area into elementary rectangles by lines parallel to the coordinate axes is, however, of special importance, and the form taken by the sums S and s for this case will therefore be explicitly stated.

Let the area be the rectangle R given by

$$x=a,\ x=b \text{ and } y=a',\ y=b',$$

and let $[a, x_1, x_2, \ldots, x_{m-1}, b]$ and $[a', y_1, y_2, \ldots, y_{n-1}, b']$ be divisions of the intervals (a, b) and (a', b') into m and n sub-intervals respectively; parallels to the coordinate axes through the points of division of these intervals will divide the rectangle R into mn rectangular meshes. If $h_r=(x_{r+1}-x_r)$ and $k_s=(y_{s+1}-y_s)$ the area of the mesh, $\sigma_{r,s}$ say, bounded by the lines $x=x_r$, $x=x_{r+1}$ and $y=y_s$, $y=y_{s+1}$ is h_rk_s; the diagonal $d_{r,s}$ of this mesh is $\sqrt{(h_r^2+k_s^2)}$ and the mesh $\sigma_{r,s}$ tends to zero in all its dimensions if and only if h_r and k_s each tend to zero. At the boundary of the area, when it is not a rectangle, the meshes will usually be only part of a rectangle.

If M, m and $M_{r,s}$, $m_{r,s}$ are the upper and lower bounds of $F(x, y)$ in R and in $\sigma_{r,s}$ respectively, then

$$S = \sum_{r,s} M_{r,s}\ h_r k_s, \quad s = \sum_{r,s} m_{r,s}\ h_r k_s \ \ldots\ldots\ldots\ldots\ldots(1)$$

where r and s take independently the values $0, 1, 2, \ldots, (m-1)$ and $0, 1, 2, \ldots, (n-1)$ respectively. S will tend to L and s to l when m and n tend to infinity provided that $\sqrt{(h_r^2 + k_s^2)}$ tends to zero; the *order* in which m and n tend to infinity is irrelevant.

A slight variation in the proof of the property, § 122, (β), namely,

$$0 \leqq S - S_1 < \mu(M-m)d^2$$

is needed. Take ξ so that $x_r < \xi < x_{r+1}$ and draw through ξ a parallel to the y-axis.

Each of the n rectangles $\sigma_{r,0}, \sigma_{r,1}, \ldots, \sigma_{r,n-1}$ will be divided into two rectangles, and if S_1' is the new value of S we shall have

$$0 \leqq S - S_1' < (M-m)h\sum_{s=0}^{n-1} k_s = (M-m)h(b'-a'),\ h_r < h.$$

If η is now taken so that $y_s < \eta < y_{s+1}$ and a parallel drawn through η to the x-axis S_1' will become S_1'' where

$$0 \leqq S_1' - S_1'' < (M-m)k(b-a), \quad k_s < k,$$

and therefore

$$0 \leqq S - S_1'' < (M-m)\{h(b'-a') + k(b-a)\}.$$

More generally, if D_1 is derived from D by inserting μ numbers between a and b and μ' numbers between a' and b', the sum S becoming S_1, we shall have

$$0 \leqq S - S_1 < (M-m)\{\mu h(b'-a') + \mu' k(b-a)\}$$

where $h_r < h$, $k_s < k$ for $r = 0, 1, \ldots, (m-1)$ and $s = 0, 1, \ldots, (n-1)$.

It is obviously possible to choose h and k so that

$$(M-m)\{\mu h(b'-a') + \mu' k(b-a)\} < \tfrac{1}{2}\varepsilon$$

as required (see § 103, equation (4)) for the proof of Darboux's Theorem.

Ex. Establish the result for an area A bounded by any curve C by enclosing A in a rectangle R, as in § 120, Theorem II.

The meshes of § 122 may of course be rectangles with sides parallel to the coordinate axes, but the division of the area just discussed supposes that the meshes are arranged in a particular way, namely, in such a way that if $x_r < \xi < x_{r+1}$, the line $x = \xi$ runs through a whole set of meshes of the same width $(x_{r+1} - x_r)$, and if $y_s < \eta < y_{s+1}$ the line $y = \eta$ runs through a whole set of meshes of the same height $(y_{s+1} - y_s)$.

In the more general case of § 122 the corresponding sets of meshes would usually have different widths and heights.

The rectangular mesh suggests another notation for the element of area in the double integral, namely

$$\int_A F(x, y)(dx\,dy)$$

where $(dx\,dy)$ takes the place of $d\sigma$; for the present the brackets are retained in the symbol for the element.

For polar coordinates the element of area would be $(r\,dr\,d\theta)$ and the integral of $F(r, \theta)$ would appear as (see *E.T.* p. 338)

$$\int_A F(r, \theta)(r\,dr\,d\theta).$$

124. Integrable Functions. The condition for the integrability of a bounded function follows at once from Darboux's Theorem.

Condition of Integrability. The condition that the bounded function $F(x, y)$ should be integrable over an area A is that, ε being given (as usual) there should be a positive number η such that $S - s$ will be less than ε when the diagonal of each mesh in the division of A for which S and s have been calculated is less than η. Or, $S - s$ must tend to zero when the diagonal of each mesh tends to zero.

It will be useful to state here another form of the definition of the double integral. If (ξ_r, η_s) is any point in the mesh σ_r, then $m_r \leqq F(\xi_r, \eta_s) \leqq M_r$, and therefore

$$s = \sum m_r\sigma_r \leqq \sum F(\xi_r, \eta_s)\sigma_r \leqq \sum M_r\sigma_r = S$$

so that $$\int_A F(x, y)\,d\sigma = \mathop{\mathcal{L}} \sum F(\xi_r, \eta_s)\sigma_r. \quad \text{.................(1)}$$

If the meshes are rectangular and (ξ_r, η_s) any point in $(h_r k_s)$

$$\int_A F(x, y)(dx\,dy) = \mathop{\mathcal{L}} \sum F(\xi_r, \eta_s)h_r k_s. \quad \text{............(2)}$$

In each case the limit is taken for the number of meshes tending to infinity in such a way that the diagonal of each mesh tends to zero.

The position of the point (ξ_r, η_s) in the rectangle $(h_r k_s)$ is arbitrary; it is permissible therefore to choose ξ_r so that the point chosen in each of the meshes contained between the lines

$x=x_r$ and $x=x_{r+1}$ shall have ξ_r as its abscissa. The limit given by (2) cannot be affected by this choice.

Of the functions that are integrable the first and most important class is that of continuous functions.

I. If $F(x, y)$ is a continuous function of x and y in A, then $F(x, y)$ is integrable over A.

The bounds M_r and m_r are values of $F(x, y)$ since $F(x, y)$ is continuous in A. Further, by the property of uniform continuity, the number η can be chosen so that $(M_r - m_r)$ will be less than ε/A for every value of r, and therefore

$$S - s < \frac{\varepsilon}{A}\Big(\sum_{r=1} \sigma_r\Big), \text{ that is, } < \varepsilon$$

when the diagonal of each mesh is less than η, and this is the condition for the integrability of $F(x, y)$ over the field A.

Cor. If $F(x, y)=1$, $\int_A (dx\,dy)=A$.

II. If $F(x, y)$ is discontinuous in A, but if its discontinuities are either finite in number or else, if infinite in number, all lie on a finite number of curves then $F(x, y)$ is integrable over A.

It must be remembered that $F(x, y)$ is bounded and that every curve is supposed to be rectifiable. It will be sufficient to prove the theorem for the case (Fig. 6) in which $F(x, y)$ is discontinuous at all points on the curve EF and on the part GFH of the bounding curve C.

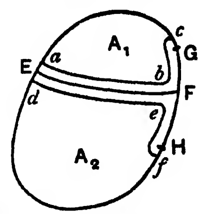

FIG. 6.

Draw curves *abc* and *def* which will cut out the lines of discontinuity from the area A; in the remaining parts, A_1 and A_2, of the area A the function $F(x, y)$ is continuous.

Now the curves *abc* and *def* may be drawn so close to EF and GFH that the area they cut out of A will be as small as we please, say less than $\varepsilon/4M$ where M is the upper limit of $F(x, y)$ in A. (Since the curves are all rectifiable, this assumption is easy to prove, if it be not considered to be "obvious.") The contribution to $S-s$ from this area is therefore less than $2M \times (\varepsilon/4M)$, that is, less than $\frac{1}{2}\varepsilon$.

The curves *abc* and *def*, when chosen as stated, are to be kept fixed. In the areas A_1 and A_2, $F(x, y)$ is continuous and

therefore, as in Theorem I, there is a division of A_1 and A_2 such that the contribution to $S-s$ from these areas is less than $\frac{1}{2}\varepsilon$. Therefore a division of the area A has been found for which $S-s<\varepsilon$ and thus $F(x, y)$ is integrable over A. (See *Note*, § 122.)

If there were other lines of discontinuity the method of proof would be the same.

Cor. If $F(x, y)$ is integrable over A the values of $F(x, y)$ may be arbitrarily changed at isolated points in A or at all points on a finite number of curves without changing the value of the integral, provided the new values of $F(x, y)$ are finite. It would be sufficient to reckon these isolated points, or the curves, among the discontinuities of $F(x, y)$; the function would still be integrable over A, and it is evident from the nature of the proof of integrability that the value of the integral would not be changed.

125. General Theorems. The following theorems are so simple that their formal proof may be left to the student. The functions $F(x, y)$, $F_1(x, y)$ and $F_2(x, y)$ are supposed to be integrable over an area A.

I. $\int_A CF(x, y)d\sigma = C\int_A F(x, y)d\sigma$, $C=$constant.

II. $\int_A \{F_1(x, y)\pm F_2(x, y)\}d\sigma = \int_A F_1(x, y)d\sigma \pm \int_A F_2(x, y)d\sigma.$

III. *The product* $F_1(x, y)\, F_2(x, y)$ *is integrable over* A.

IV. *The quotient* $F_1(x, y)/F_2(x, y)$ *is integrable over* A *if* $|F_2(x, y)| \geqq c > 0$ *in* A.

V. *When* $F(x, y)$ *is integrable over* A *so is* $|F(x, y)|$ *and*

$$\left|\int_A F(x, y)d\sigma\right| \leqq \int_A |F(x, y)|\, d\sigma.$$

VI. *If the area* A *is divided into a finite number of partial areas* $A_1, A_2, \dots,$

$$\int_A F(x, y)d\sigma = \int_{A_1} F(x, y)d\sigma + \int_{A_2} F(x, y)d\sigma + \dots.$$

VII. *Mean Value Theorem. If* $F(x, y)$ *is positive or zero in* A *then the integral of* $F(x, y)$ *is positive or zero. Hence if*

$F(x, y)=\varphi(x, y)\psi(x, y)$ *it may be deduced, as in* § 111, *that if* $\varphi(x, y) \geqq 0$, *and*

$$g \leqq \psi(x, y) \leqq G, \text{ when } (x, y) \text{ is in } A,$$

(i) $g\int_A \varphi(x, y)d\sigma \leqq \int_A \varphi(x, y)\psi(x, y)d\sigma \leqq G\int_A \varphi(x, y)d\sigma$;

(ii) $\int_A \varphi(x, y)\psi(x, y)d\sigma = K\int_A \varphi(x, y)d\sigma,\ g \leqq K \leqq G$;

(iii) *if* $\psi(x, y)$ *is continuous in* A,

$$\int_A \varphi(x, y)\psi(x, y)d\sigma = \psi(\xi, \eta)\int_A \varphi(x, y)d\sigma,\ (\xi, \eta) \text{ in } A.$$

Note. In the next article it is proved that when the discontinuities of $F(x, y)$ are of a certain type the double integral of $F(x, y)$ can be expressed as a repeated integral. It will subsequently be assumed that this restriction on $F(x, y)$ is made, unless it is explicitly stated to be removed.

126. Reduction to Repeated Integrals. It will be assumed that, if the function $F(x, y)$ is not continuous, all its discontinuities lie on a finite number of curves none of which can be cut by a line parallel to either axis in more than a finite number of points; with this restriction on the discontinuities the double integral of $F(x, y)$ exists. This restriction is not necessary, but these *admissible discontinuities* include a very wide range of functions. (See also § 127.)

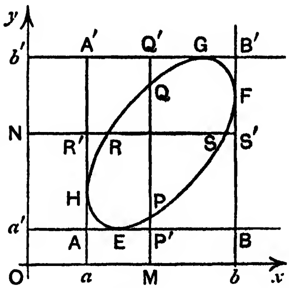

FIG. 7.

Consider first the case in which the field of integration is the rectangle $ABB'A'$, or R, given by $x=a$, $x=b$, $y=a'$, $y=b'$ (Fig. 7); then by § 124, (2),

$$\int_R F(x, y)(dx\,dy) = \underset{m,\,n}{\text{limit}} \text{ of } \sum_{r,\,s} F(\xi_r, \eta_s)h_r k_s \quad \text{.........(1)}$$

where r takes the values $0, 1, 2 \ldots, (m-1)$ and s the values $0, 1, 2, \ldots, (n-1)$ while m and n tend independently to infinity.

First, let $n \to \infty$, the numbers h_r, ξ_r, m being kept constant, and consider the sum S_n where

$$S_n = \sum_{s=0}^{n-1} F(\xi_r, \eta_s)k_s \quad .(2)$$

ξ_r is supposed to have the same value in every mesh that lies between $x=x_r$ and $x=x_{r+1}$.

The function $F(\xi_r, y)$ is a bounded function of y in (a', b') and therefore when $n\to\infty$ (and $k_s\to 0$) it has upper and lower integrals, $\varphi(\xi_r)$ and $\psi(\xi_r)$ say, where

$$\varphi(\xi_r)=\overline{\int_{a'}^{b'}}F(\xi_r, y)\,dy, \quad \psi(\xi_r)=\underline{\int_{a'}^{b'}}F(\xi_r, y)\,dy. \quad\ldots\ldots\ldots(3)$$

But, by hypothesis, $F(\xi_r, y)$ has at most a finite number of (finite) discontinuities and therefore $\varphi(\xi_r)=\psi(\xi_r)$ so that

$$\mathop{\mathcal{L}}_{n\to\infty} S_n=\varphi(\xi_r)=\psi(\xi_r)=\int_{a'}^{b'}F(\xi_r, y)\,dy. \quad\ldots\ldots\ldots\ldots(4)$$

In (1) let m now tend to infinity; the limit exists since it is equal to the double integral. Hence

$$\int_R F(x,y)(dx\,dy)=\mathop{\mathcal{L}}_{m\to\infty}\sum_{r=0}^{m-1}\varphi(\xi_r)h_r=\int_a^b\varphi(x)\,dx \quad\ldots\ldots\ldots(5)$$

so that $$\int_R F(x,y)(dx\,dy)=\int_a^b dx\int_{a'}^{b'}F(x, y)\,dy. \quad\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

Next, let m tend first to infinity, k_s, η_s and n being kept constant. As before, we find

$$\mathop{\mathcal{L}}_{m\to\infty}\sum_{r=0}^{m-1}F(\xi_r, \eta_s)h_r=\int_a^b F(x, \eta_s)\,dx$$

and $$\int_R F(x, y)(dx\,dy)=\int_{a'}^{b'}dy\int_a^b F(x, y)\,dx. \quad\ldots\ldots(7)$$

The double integral is thus expressed in two different ways by repeated integrals and the repeated integrals are equal because each is equal to the double integral.

As a corollary we have an extension of the conditions for the validity of changing the order of integration in a repeated integral with constant limits; namely, *if the discontinuities of $F(x, y)$ in the rectangle R satisfy the restriction stated at the beginning of this article, change of order of integration is permissible*, that is,

$$\int_{a'}^{b'}dy\int_a^b F(x, y)\,dx=\int_a^b dx\int_{a'}^{b'}F(x, y)\,dy.$$

Consider next the case in which the field of integration is the area A bounded by the curve C or $EFGH$ (Fig. 7); it is

supposed for the present that C cannot be cut by a parallel to either axis in more than two points. This case can be reduced to that of the rectangle R given by $x=a$, $x=b$, $y=a'$, $y=b'$.

Let $F_1(x, y)=F(x, y)$ when the point (x, y) is on or inside C but $F_1(x, y)=0$ when the point (x, y) is in R but not on or inside C. The lines AB, $A'B'$ and AA', BB' touch the curve C at E, G and H, F respectively.

The equations of HEF and HGF may be taken to be

$$y=MP=\varphi_1(x) \text{ and } y=MQ=\varphi_2(x)$$

respectively where $x=OM$; $a'=MP'$, $b'=MQ'$.

Now $$\int_R F_1(x, y)(dx\,dy)=\int_a^b dx\int_{MP'}^{MQ'} F_1(x, y)\,dy$$

since the investigation for the rectangle R remains valid provided the curve C is considered to be a curve of discontinuity for $F_1(x, y)$; the curve C satisfies the condition for a curve of discontinuity. But $F_1(x, y)=0$ for all points of R that are outside the curve C so that

$$\int_R F_1(x, y)(dx\,dy)=\int_A F(x, y)\,(dx dy),\ \int_{MP'}^{MQ'} F_1(x, y)\,dy=\int_{MP}^{MQ} F(x,y)\,dy$$

and

$$\int_A F(x, y)(dx\,dy)=\int_a^b dx\int_{\phi_1(x)}^{\phi_2(x)} F(x, y)\,dy. \quad \text{......................(8)}$$

In the same way it may be seen that if the equations of EHG and EFG are

$$x=NR=\psi_1(y) \text{ and } x=NS=\psi_2(y)$$

respectively where $y=ON$, we find

$$\int_A F(x, y)(dx\,dy)=\int_{a'}^{b'} dy\int_{\psi_1(y)}^{\psi_2(y)} F(x, y)\,dx. \quad \text{...........(9)}$$

When the area A is bounded by a curve that may be cut by a line parallel to an axis in more points than two, it may be divided into a finite number of partial areas A_1, A_2, ... bounded by curves C_1, C_2, ... each of which cannot be cut by a parallel to either axis in more than two points. The integral over A is the sum of the integrals over A_1, A_2, ... , and the reduction to repeated integrals is made for each partial area. (See Fig. 11.)

Notation. It is usual, in view of the expression in terms of repeated integrals, to denote the double integral by the double symbol $\iint$ and to omit the brackets round $dx\,dy$; thus

$$\iint_A F(x, y)dx\,dy, \quad \iint F(x, y)dx\,dy, \quad \iint F(x, y)d\sigma.$$

(See also *E.T.* § 136.)

Ex. Integrate $F(x, y)$ over the area A bounded by the circle

$$(x-\alpha)^2+(y-\beta)^2=c^2.$$

The tangents to the circle parallel to the y-axis are

$$x=\alpha-c=a \quad \text{and} \quad x=\alpha+c=b.$$

The values of y for a given x are

$$MP=\beta-\sqrt{\{c^2-(x-\alpha)^2\}} \text{ and } MQ=\beta+\sqrt{\{c^2-(x-\alpha)^2\}}$$

where the root is positive. Hence

$$I=\iint_A F(x, y)dx\,dy=\int_{\alpha-c}^{\alpha+c}dx\int_{MP}^{MQ}F(x, y)dy \quad \text{................(i)}$$

and there is a similar form if integration is first made with respect to x.

Frequently, however, it is preferable to use the polar element of area, $r\,dr\,d\theta$. In this case transfer the origin to (α, β) and then change to polar coordinates (r, θ) so that

$$x=\alpha+r\cos\theta, \quad y=\beta+r\sin\theta.$$

The limits of θ are 0 and 2π and of r are 0 and c; therefore if $F_1(r, \theta)$ is the value of $F(x, y)$ in terms of r and θ.

$$I=\iint_A F_1(r, \theta)r\,dr\,d\theta=\int_0^c r\,dr\int_0^{2\pi}F_1(r, \theta)d\theta.$$

Since the limits are constants the order of integration is easily changed.

It will be a good exercise to work out the value of the integral by both methods when

$$F(x, y)=x^2y^2\sqrt{\{c^2-(x-\alpha)^2-(y-\beta)^2\}},$$

and to verify that the value is the same for both. It should be noted that

$$\int_0^{2\pi}(A\sin^{2m}\theta\cos\theta+B\cos^{2n}\theta\sin\theta)d\theta$$

is zero when m and n are positive integers or zero; much needless labour is saved by attending to a simple matter like this.

126a. Another Proof. Let S and s be the upper and lower sums given by equation (1) of § 124 for the function $F(x, y)$, the field of integration being the rectangle R; the upper

integral L and the lower integral l of $F(x, y)$ over R exist and satisfy the inequalities

$$L \leqq S, \ l \geqq s. \quad \text{.........................(1)}$$

The only restriction on $F(x, y)$ is that it is single-valued and bounded in the rectangle R.

It will be first proved that

$$\overline{\int_a^b} dx \overline{\int_{a'}^{b'}} F(x, y)\,dy \leqq L, \quad \underline{\int_a^b} dx \underline{\int_{a'}^{b'}} F(x, y)\,dy \geqq l. \quad \text{......(2)}$$

A little consideration will show that the various steps in the proof that involve upper and lower integrals are legitimate.

Darboux's Theorem, § 122, shows that, given ε as usual, it is possible to choose d so that when the diagonal of each mesh is less than d we shall have

$$L \leqq S = \sum_{r,s} M_{r,s}\, h_r k_s < L + \varepsilon. \quad \text{.................(3)}$$

If x is fixed, say $x = \xi_r$ where $x_r \leqq \xi_r \leqq x_{r+1}$, the function $F(\xi_r, y)$ of y has an upper integral, $\varphi(\xi_r)$ say, and

$$\varphi(\xi_r) = \overline{\int_{a'}^{b'}} F(\xi_r, y)\,dy \leqq \sum_s M_{r,s}\, k_s.$$

Again, $\varphi(x)$ is a function of x which has an upper integral over (a, b) and, if $\overline{M}_r$ is the upper bound of $\varphi(x)$ in (x_r, x_{r+1})

$$\overline{\int_a^b} \varphi(x)\,dx \leqq \sum_r \overline{M}_r h_r \leqq \sum_{r,s} M_{r,s}\, h_r k_s < L + \varepsilon.$$

But ε is arbitrarily small and therefore

$$\overline{\int_a^b} \varphi(x)\,dx = \overline{\int_a^b} dx \overline{\int_{a'}^{b'}} F(x, y)\,dy \leqq L.$$

Let it be noted that the inequalities for the lower integral corresponding to those in (3) for the upper integral are

$$l - \varepsilon < s = \sum_{r,s} m_{r,s}\, h_r k_s \leqq l,$$

and it may be proved in the same way that

$$l \leqq \underline{\int_a^b} dx \underline{\int_{a'}^{b'}} F(x, y)\,dy.$$

Thus the relations (2) are established.

Suppose now that $F(x, y)$ has a double integral over R; in this case $L = l$ and the repeated integrals in (2) will therefore

also be equal to each other and to the double integral of $F(x, y)$. Hence if

$$\overline{\int_{a'}^{b'}} F(x, y)\,dy = \underline{\int_{a'}^{b'}} F(x, y)\,dy, \quad \ldots\ldots\ldots\ldots\ldots\ldots(4)$$

and therefore

$$= \int_{a'}^{b'} F(x, y)\,dy,$$

we have

$$\int_R F(x, y)(dx\,dy) = \overline{\int_a^b} dx \int_{a'}^{b'} F(x, y)\,dy = \underline{\int_a^b} dx \int_{a'}^{b'} F(x, y)\,dy,$$

that is,

$$\int_R F(x, y)(dx\,dy) = \int_a^b dx \int_{a'}^{b'} F(x, y)\,dy. \quad \ldots\ldots\ldots\ldots\ldots\ldots(5)$$

If, however, the two integrals in (4) are not equal the repeated integral

$$\int_a^b dx\, \overline{\underline{\int_{a'}^{b'}}} F(x, y)\,dy \quad \ldots\ldots\ldots\ldots\ldots\ldots(6)$$

is equal to the integral of $F(x, y)$ over R whether the upper or the lower integral with respect to y be taken; equation (5) will therefore hold even in this case provided (5) is interpreted by (6).

A similar investigation shows that equation (5) holds when the order of integration is changed. The general theorem when the field of integration is not a rectangle is dealt with as before.

Note. The student should, before reading the following articles, work through the Examples 1-6 of § 130.

127. Conditions for Repeated Integrals. There is one extension of the conditions prescribed in Article 126 that may be noticed. If one of the curves of discontinuity were a straight line parallel to a coordinate axis the double integral would still exist but there is a peculiarity as shown by the following simple example.

Ex. For the rectangle R given by $x=0, x=1, y=0, y=1$, let $F(x, y)=1$ except when $x=\frac{1}{2}$, and let $F(\frac{1}{2}, y)=+1$ for irrational values of y but $F(\frac{1}{2}, y)=-1$ for rational values of y.

From the rectangle R cut out the rectangle given by

$$x=\tfrac{1}{2}-\varepsilon, \quad x=\tfrac{1}{2}+\varepsilon', \quad y=0, \quad y=1,$$

where ε and ε' are positive and arbitrarily small, and let R' be the area that is left. The double integral of $F(x, y)$ over R' is $(1-\varepsilon-\varepsilon')$ so that the integral over R is, by definition, equal to unity.

Again

$$\int_{R'} F(x, y)(dx\,dy) = \int_0^{\frac{1}{2}-\epsilon} dx \int_0^1 1\,.\,dy + \int_{\frac{1}{2}+\epsilon'}^1 dx \int_0^1 1\,dy = 1-\varepsilon-\varepsilon'$$

$$\int_{R'} F(x, y)(dx\,dy) = \int_0^1 dy \int_0^{\frac{1}{2}-\epsilon} 1\,dx + \int_0^1 dy \int_{\frac{1}{2}+\epsilon'}^1 1\,dx = 1-\varepsilon-\varepsilon'$$

and therefore when ε and ε' tend to zero we find in this case

$$\iint_R F(x, y)\,dx\,dy = \int_0^1 dx \int_0^1 F(x, y)\,dy = \int_0^1 dy \int_0^1 F(x, y) dx.$$

The point to be noted is that $\int_0^1 F(x, y) dy$ does not exist for the value $\frac{1}{2}$ of x; if $f(x)$ denote this integral $f(x)$ is discontinuous for $x=\frac{1}{2}$. When $x=\frac{1}{2}$ the upper integral of $F(x, y)$ is $+1$, the lower integral is -1 and their difference measures the discontinuity of $f(x)$ when $x=\frac{1}{2}$.

In general, if the lines $x=c_1$, $x=c_2, \ldots, x=c_m$ are lines of discontinuity, and if $f(x)$ is given by

$$f(x) = \int_{a'}^{b'} F(x, y)\,dy,$$

$f(x)$ will be discontinuous at $c_1, c_2, \ldots, c_m$ (see § 119, Ex. 2, interchanging x and y), but $f(x)$ will still be integrable since the number of discontinuities is finite. A similar remark holds when there is a finite number of lines of discontinuity parallel to the x-axis.

It may be stated that one of the repeated integrals may exist or even that both may exist and be equal and yet the double integral not exist. (See Hobson's *Functions of a Real Variable*, 1st Ed. p. 428.) The existence of one or of both of the repeated integrals is no warrant for assuming the existence of the double integral.

128. Volume. Area of a Curved Surface. The equation of a surface, the axes of coordinates being rectangular, is in general of the form $\varphi(x, y, z)=0$, and a line parallel to the z-axis may meet the surface in more points than one; a part of the surface which is met by a line parallel to the z-axis in not more than one point will be represented by an equation of the form $z=F(x, y)$ where $F(x, y)$ is single-valued and continuous.

Suppose now that $F(x, y)$ is single-valued, continuous and positive (or at least not negative) when the point (x, y) is in an area A bounded by a closed curve C and lying in the plane $z=0$. A cylinder which has the curve C as its section by the plane $z=0$ and its generators parallel to the z-axis will intersect the surface $z=F(x, y)$ in a curve C' of which C is the projection on the xy plane; let V be the volume that is intercepted by the cylinder between its base A and the portion of the surface bounded by the curve C'. The *measure* of the volume V will now be defined.

Let the volume V be divided by two sets of planes parallel to the yz and zx planes respectively into elementary volumes that may be called *columns*; the area A will at the same time be divided into meshes that are, except possibly near the boundary, rectangular. If σ is a typical mesh and if $F(x_1, y_1)$ and $F(x_2, y_2)$ are the least and greatest values of $F(x, y)$ when (x, y) is a point in σ, the column which has σ for base will lie between two cuboids (that is, rectangular parallelepipeds) which have as their measure the products $F(x_1\, y_1)\sigma$ and $F(x_2, y_2)\sigma$. Hence the volume V will lie between two sets of cuboids whose measures are S and s where

$$S=\Sigma\, F(x_2, y_2)\,\sigma, \quad s=\Sigma\, F(x_1, y_1)\,\sigma,$$

and the summation extends over all the meshes of A.

Now $F(x, y)$ is continuous in A and therefore S and s have a common limit when the number of meshes tends to infinity and at the same time the diagonal of each mesh tends to zero; this limit is the double integral of $F(x, y)$ over A and is taken as the definition of the measure of the volume V or, when the measure of the volume is obviously meant, simply the definition of V. Hence

$$V=\int_A F(x, y)\,d\sigma=\iint_A F(x, y)\,dx\,dy. \quad \text{..............(1)}$$

Further, since the integral is independent of the shape of the meshes so long as the diagonal of each mesh tends to zero, the element of area $d\sigma$ is itself arbitrary in shape.

Again, if there are two surfaces or two parts of the same surface that are each met by a line parallel to the z-axis in only one point, their equations will be of the form $z=F(x, y)$ and $z=F_1(x, y)$, so that if $F_1(x, y)$ is greater than or equal to $F(x, y)$

the volume intercepted by the cylinder between the surfaces will be

$$\int_A F_1(x, y)d\sigma - \int_A F(x, y)d\sigma = \int_A [F_1(x, y) - F(x, y)]d\sigma.$$

The form of the result shows that the formula holds even when $F(x, y)$ is negative as would be the case, for example, if the surface were the sphere $x^2 + y^2 + z^2 = a^2$ and

$$F_1(x, y) = (a^2 - x^2 - y^2)^{\frac{1}{2}} \text{ and } F(x, y) = -(a^2 - x^2 - y^2)^{\frac{1}{2}}.$$

If the curve C is $EFGH$, Fig. 7, p. 317, then

$$V = \int_a^b dx \int_{MP}^{MQ} F(x, y)dy = \int_{a'}^{b'} dy \int_{NR}^{NS} F(x, y)dx.$$

When C may be cut by a line parallel to an axis in more points than two, it may be divided into a finite number of curves each of which will be cut by a line parallel to an axis in not more than two points and the volume would be given by the sum of the integrals over the partial areas.

Ex. 1. The volume intercepted between the plane $x + y + z = a$ and the paraboloid $2az = x^2 + y^2$ is given by the integral

$$\iint_A \left\{(a - x - y) - \frac{x^2 + y^2}{2a}\right\} dx\, dy = 4\pi a^3$$

where A is the area bounded by the circle $z = 0$, $x^2 + y^2 + 2a(x + y) = 2a^2$.

Here $F_1(x, y) = a - x - y$ and $F(x, y) = (x^2 + y^2)/2a$ and these surfaces intersect in a curve whose projection on the xy plane is the circle. For the evaluation of the integral see § 130, Ex. 6.

Area of a Curved Surface. If S is the part of the surface $z = F(x, y)$ that lies within the curve C' it is natural to assume that it has an "area," but as the surface is not in general a plane surface some definition is needed of the measure of such an area. This definition will now be given.

If $p = \partial z/\partial x = \partial F/\partial x$ and $q = \partial z/\partial y = \partial F/\partial y$, and if γ is the acute angle between the z-axis and the normal to the surface at (x, y, z)

$$\cos \gamma = |(p^2 + q^2 + 1)^{-\frac{1}{2}}|;$$

hence, if p and q are continuous, and therefore finite, when the point $(x, y, 0)$ is in the area A, the normal will not be parallel to the xy plane and therefore no tangent plane at any point of S will be perpendicular to that plane. It is assumed that p and q are continuous.

The column (in the previous construction) which has the mesh σ for its base will cut the tangent plane at any point $P(x, y, z)$ of the surface that is inside the column in a quadrilateral $t_1t_2t_3t_4$ (Fig. 8). The area, σ' say, of this quadrilateral is $\sigma \sec\gamma$ because σ is the projection of σ' on the xy plane. The sum of all the quadrilaterals for the meshes of A is

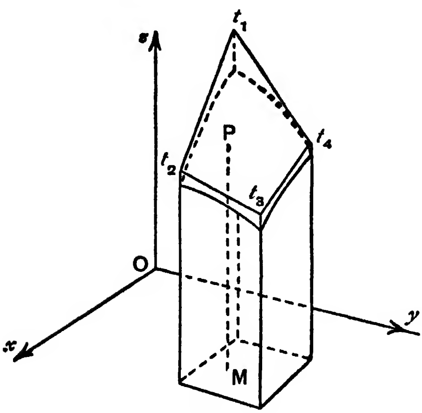

FIG. 8.

$$\Sigma \sec\gamma\,.\,\sigma = \Sigma\sqrt{(p^2+q^2+1)}\,\sigma.$$

Now $\sqrt{(p^2+q^2+1)}$ is a continuous function of x and y and therefore when the number of meshes tends to infinity, the diagonal of each mesh tending at the same time to zero, this sum has a limit, namely the integral

$$\int_A \sqrt{(p^2+q^2+1)}\,d\sigma.$$

This integral is defined to be the measure of the surface S or, as before, the definition of S when it is the measure that is clearly meant. Hence

$$S = \int_A \sqrt{\left\{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + 1\right\}}\,d\sigma. \qquad (2)$$

In the next article it is shown that this measure of the area is independent of any particular choice of the coordinate axes. See also Exercises XVI, 33.

From equation (2) it follows that $dS = d\sigma \sec\gamma$, so that, if δS is the measure of a small area of the surface at P,

$$\mathcal{L}\,\frac{\delta S}{\sigma'} = \mathcal{L}\,\frac{\delta S}{\sigma}\,.\,\mathcal{L}\,\frac{\sigma}{\sigma'} = \frac{dS}{d\sigma}\cos\gamma = 1.$$

Hence, in finding dS we may substitute σ' for δS—that is, we may suppose the area δS to be the quadrilateral $t_1t_2t_3t_4$ that lies in the tangent plane at P and the arcs that bound the area δS to be the sides of the quadrilateral. When σ is a rectangle σ' may be taken to be a parallelogram of which σ is the projection;

the area σ may, however, be of any shape, as was noted in dealing with the integral (1).

If the surface is a cylinder, $y=f(x)$, it is obvious that $dS=dz\,ds$ where ds is an element of the curve, $z=0$, $y=f(x)$.

Ex. 2. Apply the integral (2) to prove that if A, B, C are the points in which the plane

$$x/a+y/b+z/c=1$$

cuts the coordinate axes the area of the triangle ABC is

$$\tfrac{1}{2}\sqrt{(b^2c^2+c^2a^2+a^2b^2)}.$$

Here $p=-c/a$, $q=-c/b$, and therefore

$$S=\iint\sqrt{\left(1+\frac{c^2}{a^2}+\frac{c^2}{b^2}\right)}dx\,dy, \text{ taken over the triangle } OAB$$
$$=\tfrac{1}{2}\sqrt{(b^2c^2+c^2a^2+a^2b^2)}.$$

The integral thus gives the usual value, so that the two methods of measuring the area agree in this case.

Ex. 3. The area of the surface of the paraboloid $az=x^2+y^2$ that lies between the planes $z=0$ and $z=a$ is $\tfrac{1}{6}\pi(5\sqrt{5}-1)a^2$.

The projection on the xy plane of the curve in which the plane $z=a$ cuts the paraboloid is the circle (A) given by $z=0$, $x^2+y^2=a^2$; therefore

$$S=\iint_A\sqrt{\left(\frac{4x^2}{a^2}+\frac{4y^2}{a^2}+1\right)}dx\,dy.$$

Now transform to polar coordinates. (See also § 129, Ex. 5.)

129. Curves on a Surface. Element of Surface. Let the coordinates x, y, z of a point P in space, the axes being rectangular, be defined by the equations

$$x=f(u,v),\quad y=g(u,v),\quad z=h(u,v)\quad \text{.............(1)}$$

where f, g, h and their first derivatives with respect to u and v are single-valued, continuous functions when u and v vary independently within some given range. If the Jacobian $\dfrac{\partial(f,g)}{\partial(u,v)}$ is not zero u and v can be expressed as continuous functions of x and y and, when these functions are substituted for u and v in $h(u,v)$, an equation, $z=F(x,y)$ say, is obtained so that the equations (1) define a surface. It will be assumed for the present that the above Jacobian is not zero.

When v is constant, say $v=v_0$, and u varies, the equations (1) define a curve, $C(v_0)$ say, which lies on the surface and, similarly, when u is constant, $u=u_0$, and v varies, they define another curve $C(u_0)$ on the surface; the values u_0 and v_0 determine a

point P_0 on the surface which may be called "the point (u_0, v_0)" —that is, the point on the surface in which the curves $C(u_0)$ and $C(v_0)$, that is, $u=u_0$ and $v=v_0$, intersect.*

The direction cosines of the tangent at the point (u, v) to the curve $C(v)$, that is, $v=$constant, are, by § 115, Cor. 3, proportional to the derivatives f_u, g_u, h_u, and those of the tangent to the curve $C(u)$ are proportional to the derivatives f_v, g_v, h_v, so that, by the usual formulae of three dimensional coordinate geometry, if l, m, n are the direction cosines of the normal to the surface at (u, v),

$$\frac{l}{J_1}=\frac{m}{J_2}=\frac{n}{J_3}=\frac{\pm 1}{\sqrt{(EG-F^2)}}$$

where

$$\left.\begin{array}{lll} J_1=g_u h_v-h_u g_v, & J_2=h_u f_v-f_u h_v, & J_3=f_u g_v-g_u f_v \\ E=f_u^2+g_u^2+h_u^2, & F=f_u f_v+g_u g_v+h_u h_v & G=f_v^2+g_v^2+h_v^2 \end{array}\right\} \quad (2)$$

If θ is the angle between the tangents at (u, v) to $C(u)$ and $C(v)$,
$$\sqrt{(EG)}\,.\cos\theta=F, \quad \sqrt{(EG)}\,.\sin\theta=\sqrt{(EG-F^2)} \quad \text{......(3)}$$
and it is easy to prove that at the point (u, v)

$$\frac{\partial z}{\partial x}=-\frac{J_1}{J_3}, \quad \frac{\partial z}{\partial y}=-\frac{J_2}{J_3} \quad \text{......................(4)}$$

Again, an equation between u and v will define a curve C which lies on the surface. If P is a point (x, y, z) or (u, v) which lies on C and if s is the length of the arc AP, measured from any fixed point A on C, then (§ 115, Cor. 1)

$$ds^2=dx^2+dy^2+dz^2.$$

Now $dx=f_u du+f_v dv$, $dy=g_u du+g_v dv$, $dz=h_u du+h_v dv$,

and therefore $\quad ds^2=E\,du^2+2F\,du\,dv+G\,dv^2$ ....(5)

If s_1 and s_2 are the lengths of the arcs of $C(v)$ and $C(u)$ then $dv=0$ for $C(v)$ and $du=0$ for $C(u)$ so that

$$ds_1=\sqrt{E}\,du, \quad ds_2=\sqrt{G}\,dv \quad \text{.....................(6)}$$

where $\sqrt{E}$ and $\sqrt{G}$ are to be taken as *positive*. The direction cosines of the tangent to $C(v)$ are dx/ds_1, dy/ds_1, dz/ds_1; but

$$\frac{dx}{ds_1}=\frac{\partial x}{\partial u}\frac{du}{ds_1}=f_u/\sqrt{E}, \quad \frac{dy}{ds_1}=g_u/\sqrt{E}, \quad \frac{dz}{ds_1}=h_u/\sqrt{E},$$

* See Bell's *Coordinate Geometry of Three Dimensions* (2nd Ed.), pp. 348-352, with the references in the Footnote on p. 352.

with similar expressions for the direction cosines of the tangent to $C(u)$, so that as before $\sqrt{(EG)}\cos\theta = F$.

Suppose now that the curvilinear quadrilateral in Fig. 8 is that which is determined by the curves $C(u)$, $C(v)$, $C(u+du)$, $C(v+dv)$; the quadrilateral $t_1t_2t_3t_4$ in the tangent plane may, as has been seen in § 128, be substituted for δS and considered as a parallelogram with sides δs_1 and δs_2, so that the element of area dS is given by

$$dS = ds_1 ds_2 \sin\theta = \sqrt{(EG)}\,du\,dv \sin\theta = \sqrt{(EG - F^2)}du\,dv \quad \text{...(7)}$$

where $\sqrt{(EG - F^2)}$ is positive. The area S is given by the integral

$$S = \iint \sqrt{(EG - F^2)}du\,dv. \quad \text{.....................(8)}$$

Suppose next that the Jacobian J_3 or $(f_u g_v - g_u f_v)$ is identically zero. The functions $f(u, v)$ and $g(u, v)$ are therefore not independent so that f and g, or x and y, are connected by a relation, $\varphi(x, y) = 0$ say. In this case the surface is a cylinder with generators parallel to the z-axis; it may be given in general by the equations

$$x = f_1(u), \quad y = g_1(u), \quad z = h(u, v),$$

where f_1 and g_1 are functions of u alone.

Here $F = h_u h_v$, $G = h_v^2$. If we take simply $z = v$ then $F = 0$ and $G = 1$.

If a second Jacobian, J_1 say, were also zero so that an equation $\psi(y, z) = 0$ would hold in addition to $\varphi(x, y) = 0$, the equations (1) would represent a curve and not a surface.

Change of axes. If the coordinate axes are changed to another set of rectangular axes with a new origin (a, b, c), the usual equations of transformation give, ξ, η, ζ being the new coordinates,

$$x = a + l_1\xi + m_1\eta + n_1\zeta, \quad y = b + l_2\xi + \dots, \quad z = c + l_3\xi + \dots.$$

The known relations between the direction-cosines $l_1, \dots, n_3$ give

$$ds^2 = dx^2 + dy^2 + dz^2 = d\xi^2 + d\eta^2 + d\zeta^2,$$

so that the values of E, F, G are not changed and therefore the value of S, given by the integral (8), is not changed. Thus the measure of the area is independent of the coordinate axes,

as it should be. The functions E, F, G are independent of any particular choice of coordinate axes, so that the value given by (8) depends solely on the surface.

Change of parameters. If the parameters u, v are changed to u_1, v_1 by the transformation

$$u=\varphi(u_1, v_1), \quad v=\psi(u_1, v_1), \quad J=\frac{\partial(u, v)}{\partial(u_1, v_1)} \neq 0,$$

and if E_1, F_1, G_1 are the new values of E, F, G, it is not hard to show that $$\sqrt{(E_1G_1-F_1^2)}=\sqrt{(EG-F^2)}\,.\,|J|,$$
and therefore (by § 134, Problem I),

$$\iint\sqrt{(EG-F^2)}\,du\,dv=\iint\sqrt{(E_1G_1-F_1^2)}\,du_1\,dv_1.$$

Hence S is independent of the particular parameters u, v, as well as of any particular set of coordinate axes.

Ex. 1. The curves $C(u)$ and $C(v)$ are orthogonal if $F=0$.

For $\cos\theta=0$ and therefore $\theta=\pi/2$ when $F=0$. In this case dS takes the simple form $\sqrt{(EG)}\,du\,dv$.

Ex. 2. For a sphere of radius R we may put

$$x=R\sin\theta\cos\varphi, \quad y=R\sin\theta\sin\varphi, \quad z=R\cos\theta$$

and $$dS=R^2\sin\theta\,d\theta\,d\varphi$$

Here θ, φ take the place of u, v and

$$E=R^2, \quad F=0, \quad G=R^2\sin^2\theta\,; \quad \sqrt{(EG)}=R^2\sin\theta.$$

Ex. 3. For a surface of revolution about the z-axis we may put

$$x=u\cos v, \quad y=u\sin v, \quad z=F(u)$$

and $$dS=\sqrt{\{1+[F'(u)]^2\}}\,u\,du\,dv.$$

Ex. 4. If the curve given by the polar equation $r=f(\theta)$ makes a complete revolution about the initial line

$$dS=\sqrt{\{[f(\theta)]^2+[f'(\theta)]^2\}}\,r\sin\theta\,d\theta\,d\varphi.$$

Here we may take

$$x=f(\theta)\cos\theta, \quad y=f(\theta)\sin\theta\cos\varphi, \quad z=f(\theta)\sin\theta\sin\varphi.$$

Ex. 5. For the paraboloid of § 128, Example 3, we may put

$$x=u\cos v, \quad y=u\sin v, \quad z=u^2/a$$

and $$E=(4u^2+a^2)/a^2, \quad F=0, \quad G=u^2.$$

Then $$S=\iint\frac{u\sqrt{(4u^2+a^2)}}{a}\,du\,dv=\int_0^{2\pi}dv\int_0^a\frac{u\sqrt{(4u^2+a^2)}}{a}\,du,$$

so that $$S=\frac{\pi}{6}(5\sqrt{5}-1)a^2.$$

Ex. 6. If $z \equiv 0$ so that the surface is a plane surface, show that the area enclosed by a plane curve C is given by

$$\left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\, dv.$$

In this case $EG - F^2 = \dfrac{\partial(x, y)}{\partial(u, v)}$ See § 134, Problem I, Cor.

130. Worked Examples. Some examples will now be worked out to illustrate certain elements in the evaluation of double integrals.

When a double integral is given the first consideration is to determine the field of integration, and the student is strongly recommended to sketch, roughly it may be, the area over which the integration extends. There is no necessity for a detailed drawing, but the essential elements of the figure should be noted.

Ex. 1. Evaluate $\iint xy\, dx\, dy$, the field being the positive quadrant (that is, the quadrant in which both x and y are positive) of the circle $x^2 + y^2 = a^2$.

A figure shows at once that the integral is

$$\int_0^a dx \int_0^{\sqrt{(a^2 - x^2)}} xy\, dy = \int_0^a x\, dx \int_0^{\sqrt{(a^2 - x^2)}} y\, dy = \int_0^a \frac{x(a^2 - x^2)}{2}\, dx = \frac{a^4}{8}$$

Ex. 2. Evaluate $\iint y\, dx\, dy$ over the part of the plane bounded by the line $y = x$ and the parabola $y = 4x - x^2$.

The line and the parabola intersect at the points (0, 0) and (3, 3); the field of integration is that segment of the parabola that lies above the line, and if the ordinate MP at the point $P(x, y)$ on the parabola meets the line at Q the limits for the integration with respect to y are $MQ = x$ and $MP = 4x - x^2$ while the limits for the x-integration are 0 and 3. Thus

$$\iint y\, dx\, dy = \int_0^3 dx \int_x^{4x - x^2} y\, dy = \tfrac{54}{5} = 10 \cdot 8.$$

Ex. 3. Integrate $(x^2 + y^2)$ over the circle $x^2 + y^2 = a^2$.

In this case rectangular coordinates are laborious and it is simpler to suppose that the area is divided by the use of polar coordinates r, θ. The element of area is then (*E.T.* p. 338) $r\, dr\, d\theta$ and the integrand is r^2; the angle θ will vary from 0 to 2π and r from 0 to a. The integral is therefore

$$\int_0^{2\pi} d\theta \int_0^a r^3 dr = \frac{\pi a^4}{2}.$$

Ex. 4. Integrate $F(x, y)$ over the positive quadrant of the ellipse $x^2/a^2 + y^2/b^2 = 1$.

This example is taken to illustrate the method of changing the variables of integration in a double integral; at each stage the diagram should be drawn. The integral is

$$\int_0^a dx \int_0^{y_1} F(x, y)dy, \quad y_1 = \frac{b}{a}\sqrt{(a^2 - x^2)}.$$

Now let $y = b\eta$, $dy = b\,d\eta$. The limits for η are 0 and $\eta_1 = \frac{1}{a}\sqrt{(a^2 - x^2)}$,

so that $$\int_0^a dx \int_0^{y_1} F(x, y)dy = b\int_0^a dx \int_0^{\eta_1} F(x, b\eta)d\eta.$$

Let *the order of integration* be changed. The limit η_1 is an ordinate (in the first quadrant) of the ellipse $a^2\eta^2 = a^2 - x^2$. When the order of integration is changed the limits for x will be $x = 0$ and $x = x_1 = a\sqrt{(1 - \eta^2)}$ while the limits for η will be $\eta = 0$ and $\eta = 1$. Thus the integral becomes

$$b\int_0^1 d\eta \int_0^{x_1} F(x, b\eta)d\eta.$$

If now $x = a\xi$, $dx = ad\xi$ and the limits for ξ are 0 and $\xi_1 = \sqrt{(1 - \eta^2)}$ so that the integral becomes

$$ab\int_0^1 d\eta \int_0^{\xi_1} F(a\xi, b\eta)d\xi, \quad \xi_1 = \sqrt{(1 - \eta^2)}.$$

The field of integration is now the positive quadrant of the circle $\xi^2 + \eta^2 = 1$. This transformation is often useful, as in the next example.

Ex. 5. Evaluate $\iint \frac{\sqrt{(a^2b^2 - b^2x^2 - a^2y^2)}}{\sqrt{(a^2b^2 + b^2x^2 + a^2y^2)}}dx\,dy$, the area of integration being the positive quadrant of the ellipse $x^2/a^2 + y^2/b^2 = 1$.

Integral $= ab\iint \frac{\sqrt{(1 - \xi^2 - \eta^2)}}{\sqrt{(1 + \xi^2 + \eta^2)}}d\xi\,d\eta$ over pos. quad. of circle $\xi^2 + \eta^2 = 1$

$$= \frac{\pi}{2}\left(\frac{\pi}{4} - \frac{1}{2}\right)ab.$$

To evaluate the integral in ξ, η, use polar coordinates; the integral becomes, if the factor ab be omitted,

$$\int_0^1 \frac{\sqrt{(1 - r^2)}r\,dr}{\sqrt{(1 + r^2)}}\int_0^{\frac{\pi}{2}} d\theta = \frac{\pi}{2}\int_1^{\sqrt{2}} \sqrt{(2 - \varrho^2)}d\varrho, \text{ if } \varrho^2 = 1 + r^2.$$

Ex. 6. Evaluate $\iint \{2a^2 - 2a(x + y) - (x^2 + y^2)\}\,dx\,dy$, the field of integration being the circle $x^2 + y^2 + 2a(x + y) = 2a^2$.

Transfer to $(-a, -a)$ as origin by putting $x + a = \xi$, $y + a = \eta$; the integral becomes

$$\iint (4a^2 - \xi^2 - \eta^2)d\xi\,d\eta \text{ over the circle } \xi^2 + \eta^2 = 4a^2.$$

Now change to polar coordinates; the result is $8\pi a^4$.

Ex. 7. The volume common to the sphere $x^2+y^2+z^2=a^2$ and the cylinder $x^2+y^2=ay$ is $\frac{2}{9}(3\pi-4)a^3$.

The volume required is double the volume that lies above the plane $z=0$ and is therefore if $z=\sqrt{(a^2-x^2-y^2)}$,

$$2\iint z\,dx\,dy, \text{ taken over the circle } x^2+y^2=ay.$$

Transform to polar coordinates and integrate over *half* the circle, that is from $\theta=0$ to $\theta=\pi/2$; the volume is therefore

$$2\times 2\int_0^{\pi/2} d\theta\int_0^{a\sin\theta}\sqrt{(a^2-r^2)}\,r\,dr=\tfrac{2}{9}(3\pi-4)a^3.$$

Ex. 8. The volume common to the surfaces $y^2+z^2=4ax$ and $x^2+y^2=2ax$ is $\frac{2}{3}(3\pi+8)a^3$.

The volume is given by the integral

$$4\int_0^{2a}dx\int_0^{y_1}\sqrt{(4ax-y^2)}\,dy,\quad y_1=\sqrt{(2ax-x^2)}.$$

Now, *x is constant* when intégration is made with respect to y; we may change from y to θ where $y=\sqrt{(4ax)}\,.\sin\theta$ and then the integral becomes

$$4\int_0^{2a}dx\int_0^{\alpha}4ax\cos^2\theta\,d\theta,\quad \sin\alpha=\sqrt{\left(\frac{1}{2}-\frac{x}{4a}\right)}$$

$$=8a\int_0^{2a}x\left\{\sqrt{\left(\frac{1}{2}-\frac{x}{4a}\right)}\sqrt{\left(\frac{1}{2}+\frac{x}{4a}\right)}+\sin^{-1}\sqrt{\left(\frac{1}{2}-\frac{x}{4a}\right)}\right\}dx.$$

If for x is put $2a\cos\varphi$ the result comes at once.

The fact that one of the variables is constant when integration is made with respect to the other should not be forgotten when change of one of the variables is being made.

Ex. 9. When the integrand is the product of a function $\varphi(x)$ of x alone and a function $\psi(y)$ of y alone and the limits a, b for the x-integration and a', b' for the y-integration are constant, show that

$$\int_a^b dx\int_{a'}^{b'}\varphi(x)\,\psi(y)\,dy=\left(\int_a^b\varphi(x)\,dx\right)\times\left(\int_{a'}^{b'}\psi(y)\,dy\right).$$

The following examples refer to the change in the order of integration; a diagram of the field of integration is useful, and the student should make one, however rough.

Ex. 10.
$$\int_a^b dx\int_a^x F(x,y)\,dy=\int_a^b dy\int_y^b F(x,y)\,dx \quad\text{.................(i)}$$

$$\int_a^b dx\int_x^b F(x,y)\,dy=\int_a^b dy\int_a^y F(x,y)\,dx. \quad\text{..............(ii)}$$

The field in case (i) is the triangle ABC (Fig. 9), and in case (ii) the triangle ACD; AD, BC are the lines $x=a$, $x=b$ and AB, DC the lines $y=a$, $y=b$ while AC is $x=y$.

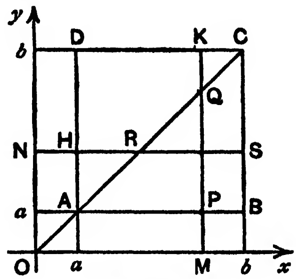

FIG. 9.

Taking the triangle ABC as the area of integration we have

$$\int_a^b dx \int_a^x F(x, y)\,dy = \int_a^b dx \int_{MP}^{MQ} F(x, y)\,dy = \int_a^b dy \int_{NR}^{NS} F(x, y)\,dx$$
$$= \int_a^b dy \int_y^b F(x, y)\,dx.$$

The other equation is proved in the same way. The special cases in which $a=0$ are frequently required and are sometimes called *Dirichlet's Formulae.*

Ex. 11. Prove that if $n \geqq 0$

$$\int_a^b dy \int_a^y (y-x)^n f(x)\,dx = \frac{1}{n+1}\int_a^b (b-x)^{n+1} f(x)\,dx.$$

Apply (ii) of Example 10; thus the double integral is equal to

$$\int_a^b dx \int_x^b (y-x)^n f(x)\,dy = \int_a^b f(x)\,.\,dx \left[\frac{(y-x)^{n+1}}{n+1}\right]_x^b = \int_a^b \frac{(b-x)^{n+1}}{n+1} f(x)\,dx.$$

Ex. 12. Prove $\displaystyle\int_0^a dx \int_0^{a-x} F(x, y)\,dy = \int_0^a dy \int_0^{a-y} F(x, y)\,dx.$

The field is the triangle bounded by $x=0$, $y=0$ and $x+y=a$.

Ex. 13. Prove that

$$\int_0^{2a} dx \int_{\frac{x^2}{4a}}^{3a-x} F(x, y)\,dy = \int_0^a dy \int_0^{2\sqrt{(ay)}} F(x, y)\,dx + \int_a^{3a} dy \int_0^{3a-y} F(x, y)\,dx.$$

(Todhunter, *Int. Cal.* p. 212.)

The field is the area bounded by the straight lines $x=0$, $y=3a-x$ and the parabola $y=x^2/4a$. The parabola and the line $y=3a-x$ intersect at A $(2a, a)$ and the field is the area OAB where $OB=3a$: the lines MN, CA, PQ are parallel to the x-axis and $OC=a$ (Fig. 10).

In this case the area OAB must be divided into the partial areas OAC and CAB; the arc OA and the line AB give $MN = 2\sqrt{(ay)}$ when

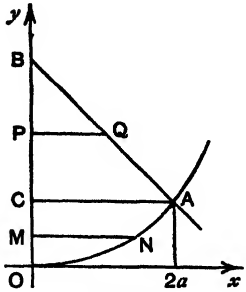

FIG. 10.

$OM = y$ and $PQ = 3a - y$ when $OP = y$. The double integral becomes, when the x-integration is taken first

$$\int_0^{OC} dy \int_0^{MN} F(x, y)dx + \int_{OC}^{OB} dy \int_0^{PQ} F(x, y)dx$$

$$= \int_0^a dy \int_0^{2\sqrt{(ay)}} F(x, y)dx + \int_a^{3a} dy \int_0^{3a-y} F(x, y)dx.$$

Ex. 14. Show that if $0 < a < b$

$$\int_a^b dx \int_{\frac{a^2}{x}}^{x} F(x, y)dy = \int_{\frac{a^2}{b}}^{a} dy \int_{\frac{a^2}{y}}^{b} F(x, y)dx + \int_a^b dy \int_y^b F(x, y)dx.$$

The field is the sector bounded by the hyperbola $xy = a^2$ and the straight lines $y = x$ and $x = b$; it must be divided into two partial areas as in Example 13.

131. Green's Theorem. Let $F(x, y)$ and $G(x, y)$ be two single-valued functions of x and y which, with the partial derivatives $\partial F/\partial y$ and $\partial G/\partial x$, are continuous when the point (x, y) is inside or on the boundary C of an area A. Green's Theorem gives the following relation between a double integral over A and a curvilinear integral along C

$$\iint_A \left(\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}\right) dx\, dy = \int_C (F dx + G dy) \quad \text{.............(1)}$$

where the integral along C is taken in the positive direction.*

* The relation (1) is a particular case of Green's general Theorem for expressing an integral taken through a volume by an integral over the surface that bounds the volume. See § 138.

Suppose first that no straight line parallel to either axis can cut C in more than two points and take the notation of Fig. 7, p. 317; then

$$\iint_A \frac{\partial G}{\partial x}\,dx\,dy = \int_{a'}^{b'} dy \int_{NR}^{NS} \frac{\partial G}{\partial x}\,dx = \int_{a'}^{b'} G(NS,\, y)\,dy - \int_{a'}^{b'} G(NR,\, y)\,dy$$

where $NS = \psi_2(y)$ and $NR = \psi_1(y)$. Hence

$$\iint_A \frac{\partial G}{\partial x}\,dx\,dy = \int_{EFG} G(x,\, y)\,dy - \int_{EHG} G(x,\, y)\,dy = \int_C G\,dy. \qquad \text{..........(2)}$$

Again,

$$\iint_A \frac{\partial F}{\partial y}\,dx\,dy = \int_a^b dx \int_{MP}^{MQ} \frac{\partial F}{\partial y}\,dy = \int_a^b F(x,\, MQ)\,dx - \int_a^b F(x,\, MP)\,dx$$

where $MQ = \varphi_2(x)$ and $MP = \varphi_1(x)$; therefore

$$\iint_A \frac{\partial F}{\partial y}\,dx\,dy = \int_{HGF} F(x,\, y)\,dx - \int_{HEF} F(x,\, y)\,dx = -\int_C F\,dx \qquad \text{........(3)}$$

From (2) and (3) equation (1) follows. The particular cases (2) and (3) should be noted. It is obvious that the curve C may consist in whole or in part of rectilinear segments; for example, C might consist of the arc FGH and the segments HA, AB, BF. Along AB, y is constant and the contribution to the integral (2) from AB is zero; similarly the contributions to the integral (3) from HA and BF are zero.

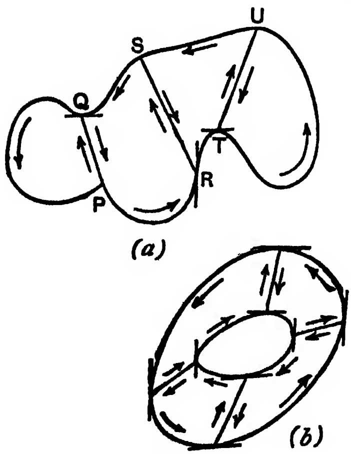

FIG. 11.

Next, if C can be cut by a line parallel to an axis in more than two points, as in Fig. 11, (a), the area may be divided by the lines PQ, RS, TU into partial areas whose boundaries satisfy the condition first imposed on C. The double integral over the whole area is the sum of the double integrals over the partial areas, while the curvilinear integrals along the auxiliary lines cancel since along each line the integral is taken twice in opposite directions. The area might be the ring-shaped region between two closed curves (Fig. 11, (b)).

Ex. 1. If $\partial G/\partial x = \partial F/\partial y$ for every point (x, y) in A and if α and β are any two points in A, prove that the integral

$$\int_{\alpha\beta}(F\,dx + G\,dy)$$

has the same value for every path from α to β, provided the path lies in A.

Let $\alpha\gamma\beta$ and $\alpha\delta\beta$ be any two curves joining α and β that lie in A and have no points in common except α and β. Green's Theorem holds for the area A' bounded by the curve $\alpha\gamma\beta\delta\alpha$ (or C'); but the double integral over A' is zero since $\partial G/\partial x = \partial F/\partial y$ at every point in A' and therefore the integral along C' is zero. Now

$$\int_{C'}(F\,dx + G\,dy) = \int_{\alpha\gamma\beta}(F\,dx + G\,dy) - \int_{\alpha\delta\beta}(F\,dx + G\,dy)$$

so that the two curvilinear integrals are equal.

Ex. 2. If the curvilinear integral in Ex. 1 is independent of the path $\alpha\beta$ when α and β are any two points in A and the path lies in A, show that $\partial G/\partial x = \partial F/\partial y$ for every point (x, y) in A.

If A' is the area bounded by *any* closed curve C' that lies in A then

$$\int_{C'}(F\,dx + G\,dy) = 0, \quad \text{so that} \quad \iint_{A'}\left(\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}\right)dx\,dy = 0.$$

Now if the continuous function $(G_x - F_y)$ is not zero at P, any point in A, there is a region A' surrounding P in which $(G_x - F_y)$ has the same sign as at P and the integral over A' could not be zero.

Ex. 3. If $\partial G/\partial x = \partial F/\partial y$ for every point (x, y) in A then $F\,dx + G\,dy$ is a complete differential (*E.T.* § 94).

Let $P(\xi, \eta)$ be any point in A. It is always possible to choose another point $M(a, b)$ in A so that the path MNP, where N is the point (ξ, b), lies in A. Let $f(\xi, \eta)$ be defined by the equation

$$f(\xi, \eta) = \int_{MNP}(F\,dx + G\,dy) + \text{const.},$$

$$= \int_a^{\xi} F(x, b)\,dx + \int_b^{\eta} G(\xi, y)\,dy + \text{const.}$$

Now $$\frac{\partial f}{\partial \xi} = F(\xi, b) + \int_b^{\eta}\frac{\partial G(\xi, y)}{\partial \xi}\,dy = F(\xi, b) + \int_b^{\eta}\frac{\partial F(\xi, y)}{\partial y}\,dy,$$

so that $$\frac{\partial f}{\partial \xi} = F(\xi, b) + F(\xi, \eta) - F(\xi, b) = F(\xi, \eta).$$

Also $$\frac{\partial f}{\partial \eta} = G(\xi, \eta).$$

Hence $$F(\xi, \eta)\,d\xi + G(\xi, \eta)\,d\eta = \frac{\partial f}{\partial \xi}d\xi + \frac{\partial f}{\partial \eta}d\eta = df(\xi, \eta),$$

or, x and y being put for ξ and η respectively,

$$F(x, y)\,dx + G(x, y)\,dy = df(x, y).$$

EXERCISES XV.

1. The volume common to the cylinders

$$x^2+y^2=a^2;\ x^2+z^2=a^2$$

$\frac{16}{3}a^3$ and the surface of one cylinder that lies inside the other is $8a^2$.

2. A sphere of radius a is pierced by a circular cylinder of radius $b\,(b<a)$, the axis of the cylinder passing through the centre of the sphere. Prove that the volume of the sphere that lies inside the cylinder is $\frac{4}{3}\pi\{a^3-(a^2-b^2)^{\frac{3}{2}}\}$ and that the surface of the sphere inside the cylinder is $4\pi a\{a-(a^2-b^2)^{\frac{1}{2}}\}$.

3. The sphere $x^2+y^2+z^2=a^2$ is pierced by the cylinder $x^2+y^2=ay$; the area of the spherical surface inside the cylinder is $2(\pi-2)a^2$.

4. An arc AB of a circle of radius r subtends the angle θ at the centre O of the circle; show that the volume of the sector of the sphere formed by the revolution about OA of the sector OAB of the circle is $\frac{2\pi}{3}r^3(1-\cos\theta)$.

Deduce the expression (*E.T.* p. 346) for the polar element of volume of a surface of revolution about the initial line.

5. One loop of the curve $r^2\cos^2\theta=a^2\cos 2\theta$ makes a complete revolution about the initial line; the volume of the solid generated is $\frac{1}{6}\pi(10-3\pi)a^3$.

6. The area of the surface of the sphere $x^2+y^2+z^2=1$ that lies inside the cylinder $2x^2(x^2+y^2)=3(x^2-y^2)$ is

$$2\pi-4\sqrt{2}\,.\,\{\sqrt{3}\log(\sqrt{3}+\sqrt{2})-2\log(1+\sqrt{2})\}.$$

7. The volume and the surface of that part of the cylinder

$$x^2/a^2+z^2/c^2=1,\ a^2-c^2=e^2a^2,$$

which lies between the planes $y=0$ and $y=mx\,(m>0)$ are $\frac{2}{3}ma^2c$ and

$$^2\left\{2+\frac{1-e^2}{e}\log\left(\frac{1+e}{1-e}\right)\right\}\text{ respectively.}$$

8. ABC is a spherical triangle and the angle ABC is $\pi/2$; if the radius of the sphere is unity, show that the area of the triangle ABC is $A+C-\frac{1}{2}\pi$ where A and C are the numbers of radians in the angles BAC and BCA.

9. The sphere $x^2+y^2+z^2=a^2$ is pierced by the cylinder

$$(x^2+y^2)^2=a^2(x^2-y^2);$$

prove that the volume of the sphere inside the cylinder is

$$\frac{8}{3}\left(\frac{\pi}{4}+\frac{5}{3}-\frac{4\sqrt{2}}{3}\right)a^3$$

and that the area of the spherical surface inside the cylinder is

$$8\left(\frac{\pi}{4}+1-\sqrt{2}\right)a^2.$$

10. The part of the volume of a sphere of radius a that lies inside a right circular cone of semi-vertical angle α whose vertex is on the sphere and whose axis is a diameter of the sphere is $\frac{4}{3}\pi(1-\cos^4\alpha)a^3$.

11. O is the centre of the ellipse $x^2/a^2+y^2/b^2=1$ and P, Q are two points on the ellipse whose eccentric angles are α, β $(\alpha<\beta)$ respectively. A cylinder with OPQ as base and with generators parallel to the z-axis intersects the paraboloid $x^2/a+y^2/b=2z$; prove that the area of the surface of the paraboloid inside the cylinder is

$$\tfrac{1}{3}(2\sqrt{2}-1)(\beta-\alpha)ab.$$

12. The area of the paraboloid $x^2/a+y^2/b=2z$ inside the cylinder $x^2/a^2+y^2/b^2=k$ is $\frac{2}{3}\pi\{(1+k)^{\frac{3}{2}}-1\}ab$.

13.* The area of the surface of the sphere $x^2+y^2+z^2=2cz$, $c>0$, inside the cone $z^2=x^2\tan^2\alpha+y^2\tan^2\beta$ is $4\pi c^2\cos\alpha\cos\beta$.

Take the coordinates $x=c\sin\theta\cos\varphi$, $y=c\sin\theta\sin\varphi$, $z=c+c\cos\theta$; the element of spherical surface is $c^2\sin\theta\,d\theta\,d\varphi$ and the values of θ and φ at the intersection of sphere and cone are connected by the equation

$$(1+\cos\theta)=(1-\cos\theta)(\cos^2\varphi\tan^2\alpha+\sin^2\varphi\tan^2\beta),$$

$$S=c^2\int_0^{2\pi}(1-\cos\theta)d\varphi=c^2\int_0^{2\pi}\frac{2\,d\varphi}{1+\cos^2\varphi\tan^2\alpha+\sin^2\varphi\tan^2\beta},\ \text{etc.}$$

14. If in Example 13 the cone is replaced by the paraboloid

$$az=x^2\tan^2\alpha+y^2\tan^2\beta$$

the area is $2\pi ac\cot\alpha\cot\beta$.

15. The area of the surface $az=xy$ that lies inside the cylinder $(x^2+y^2)^2=2a^2xy$ is $\frac{1}{6}(20-3\pi)a^2$.

16. If p is the perpendicular from the centre of the ellipsoid

$$x^2/a^2+y^2/b^2+z^2/c^2=1,\ a^2>b^2>c^2,$$

on the tangent plane at the point $P(x, y, z)$ and dS the element of area at the point, prove that, p being positive,

(i) $\displaystyle\int p\,dS=4\pi abc$;

(ii) $\displaystyle\int\frac{1}{p}\,dS=\frac{4\pi}{3abc}(b^2c^2+c^2a^2+a^2b^2)$,

the integration being in both cases over the whole surface.

Deduce from (i) the volume of the ellipsoid.

17. The surface of the ellipsoid of Example 16 is given by the integral

$$2\iint\left\{1-\frac{a^2-c^2}{a^4}x^2-\frac{b^2-c^2}{b^4}y^2\right\}^{\frac{1}{2}}\left\{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}\right\}^{-\frac{1}{2}}dx\,dy,$$

the integration being over the ellipse $x^2/a^2+y^2/b^2=1$.

Evaluate the integral for the spheroids given by (i) $b=a$, (ii) $c=b$ (*E.T.* p. 310). See also Examples 21, 22.

* Examples 13 and 14 are modified forms of examples given by Schlömilch, *Übungsbuch*, ii. pp. 281, 282.

18. The coordinates of the point (x, y, z) on the ellipsoid of Example 16 are given as

$$x = a \sin\theta \cos\varphi,\ y = b \sin\theta \sin\varphi,\ z = c\cos\theta;$$

show that dS may be expressed in the form

$$dS = \{b^2c^2 \sin^2\theta \cos^2\varphi + c^2a^2 \sin^2\theta \sin^2\varphi + a^2b^2 \cos^2\theta\}^{\frac{1}{2}} \sin\theta \, d\theta \, d\varphi,$$

and that the whole area of the surface is obtained by integrating with respect to θ and φ from 0 to π and from 0 to 2π respectively.

Evaluate for the cases $b = a$ and $c = b$.

19. If r, θ, φ are the spherical polar coordinates of a point P on a surface and γ the acute angle between the radius vector OP and the normal to the surface at P, prove that the element of surface may be expressed by the equation

$$dS = r^2 \sec\gamma \sin\theta \, d\theta \, d\varphi.$$

[Let dS' be the element of surface of a sphere with centre at the origin O and unit radius, and let a cone with O as vertex and dS' as base cut the given surface in the element dS; then $dS \cos\gamma = r^2 dS'$ so that

$$dS = r^2 \sec\gamma \, dS' = r^2 \sec\gamma \sin\theta \, d\theta \, d\varphi.]$$

20. Find, by applying the form of dS given in Example 19, the total area of the surface given by the equation

$$(x^2 + y^2 + z^2)^2 = a^2x^2 + b^2y^2 + c^2z^2,$$

and show that it is equal to the total surface of an ellipsoid whose semi-axes are bc/a, ca/b and ab/c respectively.

21. In the notation of Example 16, if θ is the angle between the normal at P and the z-axis, $\cos\theta = pz/c^2$; show that if θ is constant P lies on the curve C in which the surface

$$\cos^2\theta(x^2/a^4 + y^2/b^4 + z^2/c^4) = z^2/c^4$$

intersects the ellipsoid. Suppose $z > 0$ and $0 \leqq \theta < \pi/2$; let S be the area of the surface bounded by C (and containing the point $(0, 0, c)$) and σ the area bounded by the projection C' of C on the plane $z = 0$. Show that the surface dS lying between the curves C and C_1 that correspond to the values θ and $\theta + d\theta$ is equal to $d\sigma \sec\theta$ where $d\sigma$ is the area between the projections C' and C_1' of C and C_1 and that

$$\sigma = \pi ab(1 - \cos^2\theta)/\sqrt{\{(1 - e_1^2 \cos^2\theta)(1 - e_2^2 \cos^2\theta)\}}$$

where $$e_1^2 = (a^2 - c^2)/a^2 \text{ and } e_2^2 = (b^2 - c^2)/b^2.$$

[Obviously the equation of C' is

$$(1 - e_1^2 \cos^2\theta)\frac{x^2}{a^2} + (1 - e_2^2 \cos^2\theta)\frac{y^2}{b^2} = 1 - \cos^2\theta,$$

and the expression for σ is the area of this ellipse.]

22. The area S within the curve C (Ex. 21) is given by the integral

$$S=\int\frac{d\sigma}{\cos\theta}=\frac{\sigma}{\cos\theta}-\int\frac{\sigma\sin\theta\,d\theta}{\cos^2\theta}$$

the initial value of θ being zero.

[To evaluate the integral, let

$$e_1\cos\theta=\sin\varphi,\quad e_2^2/e_1^2=k^2<1,\quad \varphi=\sin^{-1}e_1 \text{ when } \theta=0$$

then it may be shown that $S/\pi ab$ is equal to

$$\sqrt{(1-e_1^2)}\sqrt{(1-e_2^2)}-\frac{[(1-e_1^2)-e_2^2\cos^2\varphi]\sin\varphi}{e_1\cos\varphi\sqrt{(1-k^2\sin^2\varphi)}}$$
$$+\left(\frac{1}{e_1}-e_1\right)\int_\phi^{\sin^{-1}e_1}\frac{d\varphi}{\sqrt{(1-k^2\sin^2\varphi)}}+e_1\int_\phi^{\sin^{-1}e_1}\sqrt{(1-k^2\sin^2\varphi)}\,d\varphi.$$

The transformation is somewhat laborious, but it forms a good exercise. The whole area of the surface is found by putting $\varphi=0$ and doubling the result.]

23. Apply the method of Example 22 to Example 12.

CHAPTER XI

MULTIPLE INTEGRALS. SURFACE INTEGRALS

132. Multiple Integrals. Suppose that a single-valued, bounded function $F(x, y, z)$ of three independent variables is defined for all points (x, y, z) in a volume V, the surface of the volume being included in the region of definition; the volume may be, for example, a tetrahedron or a cuboid (that is, a rectangular parallelepiped), or an ellipsoid.

If the volume be divided in any way into n elementary volumes $v_1, v_2, \dots, v_n$ and if M, m and M_r, m_r are the upper and lower bounds respectively of $F(x, y, z)$ in V and v_r we may form the sums

$$S=\sum_{r=1}^{n} M_r v_r, \quad s=\sum_{r=1}^{n} m_r v_r \quad \text{..................(1)}$$

as was done for functions of one and two variables; as before, S and s will be called the upper and lower sums for the function $F(x, y, z)$ and the particular division $[v_1, v_2, \dots, v_n]$ of V. After the discussion of the corresponding sums for the cases of functions of one and two variables there can be no difficulty in establishing similar conclusions for this case, and we will therefore simply state, without further proof, the fundamental results and then give the definitions of the integrals.

It is assumed as before that all curves are rectifiable and that all plane areas bounded by curves are quadrable; the measure of a volume is defined in § 128.

By the diagonal d_r of an element of volume v_r is meant the upper limit of the distance between two points on the surface that bounds the element; clearly $v_r < d_r^3$. When $d_r \to 0$ the element v_r will tend to zero in all its dimensions.

The sum S has a lower limit L and the sum s an upper limit l

to which they tend when n tends to infinity in such a way that the diagonal of each element of volume tends to zero.

The limits L and l are called the upper and lower triple integrals respectively of $F(x, y, z)$ over the volume (or region, or field) V, and are denoted by the symbols

$$L = \overline{\int}_V F(x, y, z)dv, \quad l = \underline{\int}_V F(x, y, z)dv. \qquad (2)$$

When $L = l$ the common limit of S and s is called the triple integral of $F(x, y, z)$ over the volume V and is denoted by the symbol

$$\int_V F(x, y, z)dv. \qquad (3)$$

The symbol dv represents the elementary volume v_r and is called, with reference to the integral, "the element of volume."

The mode of dividing V into elementary volumes is arbitrary, so long as the elementary volume tends to zero in all its dimensions when its diagonal tends to zero. If V were a cuboid given by the equations $x = a_1$, $x = b_1$, $y = a_2$, $y = b_2$, $z = a_3$, $z = b_3$, the intervals (a_1, b_1), (a_2, b_2) and (a_3, b_3) might be divided into m, n and p sub-intervals respectively and planes drawn through the points of division parallel to the coordinate planes. The typical elementary volume would then be $(h_r k_s l_t)$ where $h_r = x_{r+1} - x_r$, $k_s = y_{s+1} - y_s$, $l_t = z_{t+1} - z_t$, the numbers x_r, y_s, z_t being representative numbers in the intervals (a_1, b_1), (a_2, b_2), (a_3, b_3). If (ξ_r, η_s, ζ_t) is any point in the elementary volume $(h_r k_s l_t)$ the triple integral over V would be defined as the limit of the sum

$$\sum_{r,s,t} F(\xi_r, \eta_s, \zeta_t) h_r k_s l_t \qquad \begin{array}{l} r = 0, 1, 2, \ldots, (m-1) \\ s = 0, 1, 2, \ldots, (n-1) \\ t = 0, 1, 2, \ldots, (p-1) \end{array} \qquad (4)$$

when m, n and p tend to infinity in such a way that the diagonal $\sqrt{(h_r^2 + k_s^2 + l_t^2)}$ of each elementary volume tends to zero. The corresponding notation for the triple integral would be

$$\int_V F(x, y, z)(dx\,dy\,dz) \text{ or } \iiint_V F(x, y, z)(dx\,dy\,dz) \qquad (5)$$

the element of volume being now $(dx\,dy\,dz)$. The three symbols $\int$ of integration become appropriate when the evaluation of the triple integral is made by three repeated integrations; it is usual then to omit the brackets round $dx\,dy\,dz$.

Integrable Functions. The condition that $F(x, y, z)$ should be integrable over V is found as before; L will be equal to l if there is a division of V for which $S-s$ is less than ε (see § 122, Note), or if $S-s$ tends to zero when the diagonal of each elementary volume tends to zero.

If $F(x, y, z)$ is continuous in V it may be shown as before, by using the property of uniform continuity, that $F(x, y, z)$ is integrable over V. If $F(x, y, z)$ is discontinuous in V it will be integrable over V if its discontinuities are finite in number or if, when infinite in number, they all lie on a finite number of surfaces, which can therefore be enclosed in a finite number of volumes whose total volume can be made arbitrarily small.

The reduction to repeated integrals is effected by the same method as in § 126. Take the case given by (4) for the cuboid. In the sum keep all the numbers except those that refer to the division of (a_3, b_3) constant and let S_p be the sum

$$S_p = \sum_{t=0}^{p-1} F(\xi_r, \eta_s, \zeta_t) l_t;$$

when $p \to \infty$ this sum tends to a limit, $\varphi(\xi_r, \eta_s)$ say, where

$$\varphi(\xi_r, \eta_s) = \int_{a_3}^{b_3} F(\xi_r, \eta_s, z) dz,$$

because $F(\xi_r, \eta_s, z)$ is either continuous or has only a finite number of discontinuities. The sum $\Sigma\varphi(\xi_r, \eta_s) h_r k_s$ can now be treated in the same way and we find

$$\iiint_V F(x, y, z)\, dx\, dy\, dz = \int_{a_1}^{b_1} dx \int_{a_2}^{b_2} dy \int_{a_3}^{b_3} F(x, y, z) dz. \quad \ldots(6)$$

The order in which m, n, p are made to tend to infinity makes no difference to the value of the triple integral and therefore the six repeated integrals, of which that in equation (6) is one, are all equal; in other words the order of integration, when the limits are all constants, is indifferent, just as for the case of two variables. The repeated integrals will exist even though there be a finite number of planes of discontinuity parallel to the coordinate axes though the separate integrals may not exist (see § 127).

If V is not a cuboid the reduction to repeated integrals may be effected in the same way as for a double integral by enclosing

V in a cuboid and taking $F_1(x, y, z) = F(x, y, z)$ for points in V, but $F_1(x, y, z) = 0$ for points in the cuboid that are not in V. (See § 126.) See also *E.T.* pp. 338, 339.

If $F(x, y, z, w)$ is defined for a four-dimensional region R, say for the points (x, y, z, w) where

$$a_1 \leqq x \leqq b_1,\ a_2 \leqq y \leqq b_2,\ a_3 \leqq z \leqq b_3 \text{ and } a_4 \leqq w \leqq b_4,$$

the quadruple integral

$$\int_R F(x, y, z, w)(dx\,dy\,dz\,dw)$$

would be defined as the common limit of sums S and s where

$$S = \sum_{r=1}^{n} M_r W_r \quad s = \sum_{r=1}^{n} m_r W_r$$

and M_r, m_r are the upper and lower limits of the function $F(x, y, z, w)$ in the elementary region W_r, and the integral would be reducible to four repeated integrals

$$\int_R F(x, y, z, w)\,dx\,dy\,dz\,dw = \int_{a_1}^{b_1} dx \int_{a_2}^{b_2} dy \int_{a_3}^{b_3} dz \int_{a_4}^{b_4} F(x, y, z, w)\,dw.$$

In the same way quintuple, sextuple, ... , n-ple integrals may be defined.

Ex. 1. $\iiint \dfrac{dx\,dy\,dz}{(x+y+z+1)^3}$ throughout the volume bounded by the planes $x=0$, $y=0$, $z=0$, $x+y+z=1$.

See Fig. 78 (*E.T.* p. 339) and let $OA = OB = OC = 1$. The integral is

$$\int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} \frac{dz}{(x+y+z+1)^3} = \tfrac{1}{2}\log 2 - \tfrac{5}{16}.$$

Ex. 2. If $x^2+y^2+z^2=r^2$, calculate the integral of r^2 when the field of integration is the volume inside the sphere $x^2+y^2+z^2=a^2$.

In this case it is most convenient to take the polar element of volume $r^2 \sin\theta\,dr\,d\theta\,d\varphi$ (*E.T.* p. 346). The integral is

$$\int_0^a r^4 dr \int_0^\pi \sin\theta\,d\theta \int_0^{2\pi} d\varphi = \frac{4\pi}{5}a^5.$$

Examples 6-14 (*E.T.* pp. 348, 349) furnish easy cases of double and triple integrals.

133. Change of Variables. The method will first be considered for special cases that will illustrate the general process and emphasize important details.

The first point that must be grasped is that in the case of a double integral, for example, when the variables x and y are changed to u and v by the substitutions $x = f(u, v)$, $y = g(u, v)$

and u is to take the place of x and v the place of y, *the change from y to v is made on the supposition that x is constant;* the variable u may be supposed to be eliminated between the equations $x=f(u, v)$, $y=g(u, v)$ so as to give an equation $y=\varphi(x, v)$. If u is not eliminated the relation between dy and dv would be obtained by differentiating the equations $x=f$, $y=g$, treating x as constant and eliminating du from the equations

$$0=\frac{\partial f}{\partial u}du+\frac{\partial f}{\partial v}dv, \quad dy=\frac{\partial g}{\partial u}du+\frac{\partial g}{\partial v}dv.$$

Next when v has taken the place of y the order of integration must be changed; the value of dx is then obtained from the equation between x, u and v, the variable v being now treated as a constant.

Care must be taken to assign the limits of the new integrals properly, and in any given example the use of a diagram is strongly recommended.

Ex. 1. Transform the integral $I=\int_0^a dx\int_0^{a-x}F(x, y)\,dy$ by the substitutions $x+y=u$, $x=uv$.

We can at once express y in terms of x and u and since the first integration is with respect to y we replace y by u; thus we find

$$I=\int_0^a dx\int_x^a F(x, u-x)\,du=\int_0^a dx\int_x^a F_1(x, u)\,du \quad\text{.................(i)}$$

where $F_1(x, u)$ is the value of $F(x, y)$ in terms of x and u.

The next step is to change the order of integration; we get

$$I=\int_0^a du\int_0^u F_1(x, u)\,dx. \quad\text{.............................(ii)}$$

Finally from the equation $x=uv$ we have $dx=udv$ and the limits for v are 0 and 1 so that

$$I=\int_0^a du\int_0^1 F_1(uv, u)\,udv=\int_0^a du\int_0^1 F_2(u, v)\,udv$$

where $F_2(u, v)$ is the value of $F(x, y)$ in terms of u and v.

The integral has therefore been expressed as an integral with constant limits.

Ex. 2. By means of the substitution in Example 1, show that

$$\int_0^1 dx\int_0^{1-x} x^{m-1}y^{n-1}(1-x-y)^{p-1}dy=\int_0^1 u^{m+n-1}(1-u)^{p-1}du\int_0^1 v^{m-1}(1-v)^{n-1}dv$$

where m, n, p are each not less than unity (so that the integrand may be bounded). Prove that the value of the integral is

$$\Gamma(m)\Gamma(n)\Gamma(p)/\Gamma(m+n+p).$$

Ex. 3. Transform the integral $I=\int_0^a dx\int_0^{a-x}dy\int_0^{a-x-y}F(x, y, z)dz$ by the substitution $\quad x+y+z=u,\ x+y=uv,\ x=uvw.$

First, let $z=u-x-y$ and change from z to u; then

$$I=\int_0^a dx\int_0^{a-x}dy\int_{y+x}^a F_1(x, y, u)du, \text{ ....(i)}$$

where $F_1(x, y, u)$ is the value of $F(x, y, z)$ in terms of x, y, u.

The next step is to change the order of integration. The integral

$$\int_0^{a-x}dy\int_{y+x}^a F_1(x, y, u)du$$

is taken over the triangle bounded by $y=0$, $u=a$, $u=y+x$ (x constant); a diagram will show at once that the limits for y are 0 and $u-x$ and for u are x and a. Hence

$$I=\int_0^a dx\int_x^a du\int_0^{u-x}F_1(x, y, u)dy=\int_0^a dx\int_x^a \varphi(x, u)du, \text{ say}$$

where
$$\varphi(x, u)=\int_0^{u-x}F_1(x, y, u)dy.$$

The change of order in this value of I is given by Ex. 10, § 130, so that

$$I=\int_0^a du\int_0^u dx\int_0^{u-x}F_1(x, y, u)dy. \text{ ....(ii)}$$

Now take the equation $y=uv-x$; keep u and x in (ii) constant and then $dy=u\,dv$ so that we find

$$I=\int_0^a du\int_0^u dx\int_{\frac{x}{u}}^1 F_2(x, v, u)u\,dv \text{ ....(iii)}$$

where $F_2(x, v, u)$ is the value of $F(x, y, z)$ in terms of x, u, v.

Again, the order of integration with respect to x and v has to be changed; by Ex. 10, § 130, slightly modified, the integral becomes

$$I=\int_0^a du\int_0^1 dv\int_0^{uv}F_2(x, v, u)u\,dx \text{ ....(iv)}$$

Finally, let $x=uvw$ and then $dx=uv\,dw$, so that

$$I=\int_0^a du\int_0^1 dv\int_0^1 F_3(w, v, u)u^2v\,dw \text{ ....(v)}$$

where $F_3(w, v, u)$ is the value of $F(x, y, z)$ in terms of u, v, w.

The new integral has constant limits.

Ex. 4. In Example 3 let

$$a=1 \quad\text{and}\quad F(x, y, z)=x^{m-1}y^{n-1}z^{p-1}(1-x-y-z)^{r-1}$$

where m, n, p, r are each not less than unity so that $F(x, y, z)$ may be bounded; then prove that the value of the integral is

$$\Gamma(m)\Gamma(n)\Gamma(p)\Gamma(r)/\Gamma(m+n+p+r).$$

Ex. 5. Prove that the integral

$$\int_0^1 dx_1 \int_0^{1-x_1} dx_2 \int_0^{1-x_1-x_2} dx_3 \int_0^{1-x_1-x_2-x_3} F(x_1, x_2, x_3, x_4)\, dx_4$$

can be expressed as a quadruple integral with the limits 0 and 1 for each of the new variables. Extend the result to an n-ple integral.

Ex. 6. If $I = \int_0^a dx \int_0^x F(x, y)\,dy$ and $x(1+u) = v$, $y = xu$, prove that

$$I = \int_0^1 du \int_0^{a(1+u)} F_2(u, v) \frac{v}{(1+u)^2}\, dv,$$

where $F_2(u, v)$ is the value of $F(x, y)$ in terms of u, v.

Here the equation $y = xu$ at once suggests that u should take the place of y and then we find

$$I = \int_0^a dx \int_0^1 F(x, xu)x\, du = \int_0^1 du \int_0^a F(x, xu)x\, dx.$$

The change of order is easy since the limits are constants. We then take the equation $x(1+u) = v$; since u is constant for the x-integration $dx = (1+u)^{-1}dv$ and the result comes at once.

Another method may be adopted for integrals of the type of Examples 3 and 4; we take $a = 1$ and transform the integral in Example 3.

Ex. 7. First, let $z = (1-x-y)\zeta$; therefore

$$\int_0^{1-x-y} F(x, y, z)\,dz = (1-x-y)\int_0^1 F[x, y, (1-x-y)\zeta]\,d\zeta$$

$$= (1-x-y)\varphi(x, y), \text{ say,}$$

because the integral is a function of x and y alone.

Next, let $y = (1-x)\eta$ so that $1-x-y = (1-x)(1-\eta)$; then

$$\int_0^{1-x} (1-x-y)\varphi(x, y)\,dy = (1-x)^2 \int_0^1 (1-\eta)\varphi[x, (1-x)\eta]\,.\,d\eta$$

where $\varphi[x, (1-x)\eta] = \int_0^1 F[x, (1-x)\eta, (1-x)(1-\eta)\zeta]\,d\zeta$.

Lastly, let $x = \xi$, for symmetry, and we find that the given integral becomes

$$\int_0^1 (1-\xi)^2 d\xi \int_0^1 (1-\eta)\,d\eta \int_0^1 F[\xi, (1-\xi)\eta, (1-\xi)(1-\eta)\zeta]\,d\zeta$$

$$= \int_0^1\int_0^1\int_0^1 F[\xi, (1-\xi)\eta, (1-\xi)(1-\eta)\zeta](1-\xi)^2(1-\eta)\, d\xi\, d\eta\, d\zeta.$$

The transformation is given by

$$x = \xi,\ y = (1-\xi)\eta,\ z = (1-\xi)(1-\eta)\zeta.$$

The method obviously applies to Example (4) and the general integral of the same type. Note that

$$\frac{\partial(x, y, z)}{\partial(\xi, \eta, \zeta)} = (1-\xi)^2(1-\eta).$$

134. General Method. In the method of the preceding article the new variables are introduced in succession, and when the first new variable has been introduced a change in the order of integration is necessary before the second new variable can be brought in. In the general method that will now be explained the same procedure will be adopted, but no attempt will be made to *specify the limits of the individual* integrals either before or after the transformation has been made. The equations that connect the old variables $x, y, z, \ldots$ with the new $u, v, w, \ldots$ will transform the old field of integration into a new field, and when the integral over the given field has been transformed into one over the new field the actual specification of the limits of the integral in both forms of it is left for determination in each particular case. The following observations may be useful.

Let I be the integral of $F(x, y, z)$ over a region A and suppose that it is expressed as a repeated integral, say

$$I = \int_a^b dx \int_{\phi_1(x)}^{\phi_2(x)} dy \int_{\psi_1(x,y)}^{\psi_2(x,y)} F(x, y, z)\, dz.$$

The integral will be said to be in *standard form* if the upper limit of each of the repeated integrals is algebraically greater than the lower limit. The student might write down the values of the functions $\varphi_1, \varphi_2, \psi_1, \psi_2$ for a region bounded by a simple surface (such as an ellipsoid) which is cut in not more than two points by a line parallel to any coordinate axis; there is no real limitation in supposing that the region can always be divided into partial regions that satisfy this condition and the integral over the whole region is the sum of the integrals over the partial regions. In any particular case this division is usually made (see, for example, § 130, Examples 13, 14).

It will be assumed that every integral is expressed in standard form and that the change of order of integration can be effected though the actual limits of the integrals will not be specified. The importance of the assumption of the standard form will be seen in the discussion of Problem I.

Again it will be assumed that the old variables are expressed in terms of the new by functions which are not only continuous

but have continuous partial derivatives of the first and second orders and that the correspondence is "one-to-one"—that is, that to each point, (x, y, z) say, within and not on the boundary of the old region there corresponds one and only one point, (u, v, w) say, within and not on the boundary of the new region into which the old region is transformed by the equations of transformation.

Problem I. Two variables. The field is a region A and the integral is I where

$$I=\iint_A F(x, y)\,dx\,dy=\int dx\int F(x, y)\,dy=\int dy\int F(x, y)\,dx.$$

Let the transformation be given by the equations

$$x=f(u, v),\quad y=g(u, v) \quad \text{.......................(1)}$$

and let J be the Jacobian

$$J=J(f, g)=\begin{vmatrix} f_u, & f_v \\ g_u, & g_v \end{vmatrix} \quad \begin{vmatrix} x_u, & x_v \\ y_u, & y_v \end{vmatrix}$$

If J is not zero and *does not change sign* the correspondence is one-to-one (§ 56, Theorem III, Ex.).

(i) Substitute for y in terms of v. In this operation x is constant and u may be considered to be a parameter; theoretically u might be eliminated between equations (1) and y expressed as a function of x and v but, though this elimination is sometimes useful, it is in general impracticable. Of course, if x is constant and v varies, u must also vary, but this variation is taken into account in finding dy. Take the differentials of the functions in equations (1); then

$$0=f_u du+f_v dv,\quad dy=g_u du+g_v dv$$

and therefore, solving for dy,

$$dy=\frac{f_u g_v-f_v g_u}{f_u}\,dv=\frac{J}{f_u}\,dv.$$

Now the integral I is in standard form and therefore the upper limit, y_2 say, of the y-integral is algebraically greater than the lower limit y_1 and dy is positive; let the values of v given by the transformation be v_1, v_2 when those of y are y_1, y_2 respectively.

If J/f_u is positive v will increase when y increases and therefore $v_2>v_1$. On the other hand, if J/f_u is negative v will decrease

as y increases and therefore $v_2 < v_1$; the upper limit of the v-integral when in standard form will now be v_1 and this interchange of the limits is made by changing the sign of the v-integral, or the sign of the integrand, so that $(J/f_u)dv$ becomes $(-J/f_u)dv$. In both cases therefore

$$dy = \qquad dv, \quad I = \int dx \int F_1(x, v) \qquad dv$$

where $F_1(x, v) = F(x, y)$—that is, $F_1(x, v)$ is the value of $F(x, y)$ in terms of x and v.

In general, therefore, however many variables of each set there may be, when the transformation gives $dy = \varphi \,.\, dv$, where φ is a function of v and other variables, the form to be substituted for dy is not $\varphi\, dv$ but $|\varphi|\, dv$. This form will now be used, without further remark, in all cases.

(ii) The next step is to change the order of integration, and this change is assumed to be made so that

$$I = \int dv \int F_1(x, v) \left|\frac{J}{f_u}\right|$$

(iii) Finally, substitute for x in terms of u, keeping v constant; then $dx = |f_u|\, du$ so that

$$I = \int dv \int F_2(u, v)\, |J|\, du = \iint F_2(u, v)\, |J|\, du\, dv$$

where $F_2(u, v) = F(x, y)$.

If f_u were identically zero the above process would fail because of the value it gives for dy. But in this case f_v cannot be identically zero because, if it were, J would be zero and x would be constant. Consequently we can begin by substituting for y in terms of u. We may, however, begin by substituting for x in terms of v and then $dx = |f_v|\, dv$, and the integral becomes

$$I = \int dy \int F_1(y, v) |f_v| dv = \int dv \int F_1(y, v) |f_v| dy.$$

Next, $dy = |g_u| du$ so that

$$I = \int dv \int F_2(u, v) |f_v| |g_u| du = \iint F_2(u, v) |J| du\, dv,$$

the same value as before, because $|J| = |f_v g_u|$.

Cor. The element of area dA in terms of the coordinates u, v is
$$dA = |J|\,du\,dv,$$
as may be seen by supposing $F(x, y) = 1$.

Ex. 1. If $x = r\cos\theta$, $y = r\sin\theta$ then $|J| = r$ so that
$$\iint F(x, y)\,dx\,dy = \iint F_1(r, \theta)\,r\,dr\,d\theta,\quad F_1(r, \theta) = F(x, y).$$
This example shows that J may be zero at isolated points, provided it does not change sign.

Ex. 2. If $x = a\xi$, $y = b\eta$, $a > 0$, $b > 0$, then $J = ab > 0$ so that
$$\iint_A F(x, y)\,dx\,dy = ab\iint_{A'} F_1(\xi, \eta)\,d\xi\,d\eta.$$
If the field A is the ellipse $x^2/a^2 + y^2/b^2 = 1$ the field A' is the circle $\xi^2 + \eta^2 = 1$; if the field A is a sector of the ellipse the field A' is the corresponding sector of the circle.

Problem II. Three variables. Let the transformation be
$$x = f(u, v, w),\quad y = g(u, v, w),\quad z = h(u, v, w). \quad \text{.........(2)}$$
The Jacobian $J = \dfrac{\partial(f, g, h)}{\partial(u, v, w)}$ must be different from zero and always of the same sign; the correspondence will then be one-to-one.

The work may be carried out in this case with less fulness of detail.

(i) Substitute for z in terms of w. The differential dz is found by taking the differentials in equations (2), x and y being constant; then
$$0 = f_u\,du + \ldots,\quad 0 = g_u\,du + \ldots,\quad dz = h_u\,du + h_v\,dv + h_w\,dw,$$
so that, if $J_3 \neq 0$, $\quad dz = \dfrac{J}{J_3}\,dw,\ J_3 = \dfrac{\partial(f, g)}{\partial(u, v)}$.

Hence
$$I = \int dx\int dy\int F_1(x, y, w)\left|\frac{J}{J_3}\right|dw,\quad F_1(x, y, w) = F(x, y, z).$$

(ii) Change the order of integration and substitute for y in terms of v; dy is found from the equations $x = f$ and $y = g$, when x and w are kept constant. Hence
$$0 = f_u\,du + f_v\,dv,\quad dy = g_u\,du + g_v\,dv,$$
so that
$$dy = (J_3/f_u)\,dv,$$
and $I = \displaystyle\int dw\int dx\int F_2(x, v, w)\left|\frac{J_3}{f_u}\right|dv,\quad F_2(x, v, w) = F(x, y, z).$

(iii) Change the order of integration and substitute for x in terms of u; then $dx = f_u\, du$ and

$$I = \int dw \int dv \int F_3(u, v, w)|J|du = \iiint F_3(u, v, w)|J|du\, dv\, dw$$

where $F_3(u, v, w)$ is the value of $F(x, y, z)$ in terms of u, v, w.

If J_3 is identically zero there must be at least one of the first minors of J, say $\partial(g, h)/\partial(v, w)$, that is not identically zero, otherwise J would be zero. The differential dx, when y and z are constant, is given by $(J/J_1)du$ where J_1 is the above minor; begin therefore by substituting for x in terms of u. In all cases the form of the resultant integral is the same.

Cor. The element of volume dV in terms of the coordinates u, v, w is

$$dV = |\, J \,|\, du\, dv\, dw,$$

as may be seen by supposing $F(x, y, z) = 1$. See also *Exercises* XVI, 34.

Ex. 3. If $x = r \sin \theta \cos \varphi,\ y = r \sin \theta \sin \varphi,\ z = r \cos \theta$ then $J = r^2 \sin \theta$ and

$$F(x, y, z)\, dx\, dy\, dz = \iiint F_1(r, \theta, \varphi)\, r^2 \sin \theta\, dr\, d\theta\, d\varphi.$$

See remark on Ex. 1.

Ex. 4. If $x = a\xi,\ y = b\eta,\ z = c\zeta,\ a > 0,\ b > 0,\ c > 0,\ J = abc$

and $$\iiint F(x, y, z)\, dx\, dy\, dz = abc \iiint F(a\xi, b\eta, c\zeta)\, d\xi\, d\eta\, d\zeta.$$

(Compare Ex. 2.)

Problem III. Implicit Functions. If the old and the new variables are connected by equations of the form

$$\varphi(x, y, z, u, v, w) = 0,\quad \psi(x, \dots, w) = 0,\quad \chi(x, \dots, w) = 0,$$

express the Jacobian J by means of the relation, § 55, (3),

$$\frac{\partial(\varphi, \psi, \chi)}{\partial(u, v, w)} = (-1)^3 \frac{\partial(\varphi, \psi, \chi)}{\partial(x, y, z)} \cdot J.$$

The transformed value of the integral may therefore be expressed by using the solution in Problem II, when x, y, z have been determined as functions of u, v, w; this is possible since obviously the Jacobians are supposed to be different from zero.

Problem IV. n variables. If there are n variables

$x_1, x_2, \dots, x_n$ connected with the new variables $y_1, y_2, \dots, y_n$ by the equations

$$x_r = f_r(y_1, y_2, \dots, y_n),\ r = 1, 2, \dots, n,$$

the solution will obviously be

$$\iint\dots\int F(x_1, x_2, \dots, x_n)\,dx_1 dx_2 \dots dx_n = \iint\dots\int F_1 |J|\, dy_1 dy_2, \dots dy_n$$

where F_1 is the value of $F(x_1, x_2, \dots, x_n)$ in terms of $y_1, y_2, \dots, y_n$ and

$$J = \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(y_1, y_2, \dots, y_n)}.$$

If the variables are given as in Problem III by equations of the form

$$\varphi_r(x_1, \dots, x_n, y_1, \dots, y_n) = 0,\ r = 1, 2, \dots, n,$$

proceed as in Problem III.

Ex. 5. If the variables x, y, z are changed to ξ, η, ζ by a properly chosen orthogonal transformation, show that

$$\iiint F(ax + by + cz)\,dx\,dy\,dz = \iiint F(k\xi)\,d\xi\,d\eta\,d\zeta \quad \dots\dots\dots\dots(1)$$

where $k = |(a^2 + b^2 + c^2)^{\frac{1}{2}}|$ and the region of integration in each case is a sphere of radius unity with centre at the origin of coordinates.

The new and the old variables are connected (Bell, *Coordinate Geometry of Three Dimensions*, Chap. IV) by equations of the form

$$\xi = l_1 x + m_1 y + n_1 z,\quad \eta = l_2 x + m_2 y + n_2 z,\quad \zeta = l_3 x + m_3 y + n_3 z$$

where $$\xi^2 + \eta^2 + \zeta^2 = x^2 + y^2 + z^2$$

and the coefficients $l_1, \dots, n_3$ satisfy certain conditions—the conditions of orthogonality. (See Bell, *l.c.*, Equations (A), (B), (C), (D) of § 53, 2nd Ed.)

Now let $a = kl_1$, $b = km_1$, $c = kn_1$; this choice is possible provided $k^2 = a^2 + b^2 + c^2$ and k will be taken to be positive.

Again,

$$\frac{\partial(\xi, \eta, \zeta)}{\partial(x, y, z)} = \begin{vmatrix} l_1, & m_1, & n_1 \\ l_2, & m_2, & n_2 \\ l_3, & m_3, & n_3 \end{vmatrix} = \pm 1 = \frac{\partial(x, y, z)}{\partial(\xi, \eta, \zeta)}.$$

Since $|J| = 1$ and $ax + by + cz = k\xi$ the equation (1) follows at once and the region of integration in the new integral is the sphere given by $\xi^2 + \eta^2 + \zeta^2 = 1$.

The theorem may obviously be extended to the case of n variables $x_1, x_2, \dots, x_n$ when the transformation to the new variables $\xi_1, \xi_2, \dots, \xi_n$ is orthogonal—that is, when the coefficients of the transformation satisfy equations corresponding to the equations (A) ... (D) mentioned above so that $$x_1^2 + x_2^2 + \dots + x_n^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_n^2.$$

The function $F(a_1x_1+a_2x_2+\ldots+a_nx_n)$ would become $F(k\xi_1)$ where k is $|(a_1^2+a_2^2+\ldots+a_n^2)^{\frac{1}{2}}|$; the Jacobian $|J|=1$ and the new variables are such that $0 \leqq \xi_1^2+\xi_2^2+\ldots+\xi_n^2 \leqq 1$, it being understood that the old variables are such that $0 \leqq x_1^2+x_2^2+\ldots+x_n^2 \leqq 1$.

The student should work some of the earlier examples in *Exercises* XVI before reading further in the text.

135. Surface Integrals. On a surface given by the equation $\varphi(x, y, z)=0$ let there be a portion S, bounded by a closed curve C, which is such that it cannot be cut by any line parallel to the coordinate axes in more than one point; the relation between the coordinates of any point on S may therefore be expressed in any one of the forms

$$x=f(y, z), \quad y=g(z, x), \quad z=h(x, y) \quad \ldots\ldots\ldots\ldots\ldots(1)$$

where f, g, h are single-valued, continuous functions of their variables.

Let C_1, C_2, C_3 be the projections of C on the coordinate planes of yz, zx, xy respectively and S_1, S_2, S_3 the areas enclosed by these curves.

If $F(x, y, z)$ is single-valued and continuous when (x, y, z) is any point in S the function depends only on two variables because one of them may be eliminated by using the equations (1). Suppose that z is eliminated so that $F(x, y, z)$ becomes $F\{x, y, h(x, y)\}$ or $F_3(x, y)$. The area S_3 in the xy plane is the projection of S and the definition is now made:

Definition. The double integral of $F_3(x, y)$ over S_3, that is,

$$\iint_{S_3} F_3(x, y)\, dx\, dy, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

is a surface integral of $F(x, y, z)$ over the surface S.

Similar definitions hold when the variables x or y are eliminated and the double integral is taken over S_1 in the plane of yz or over S_2 in the plane of zx so that there are three types of surface integrals when x, y, z are the variables.

If l, m, n are the direction cosines of the normal to the surface S at (x, y, z), and if the element dS projects into the element $dx\,dy$ then, if n is positive, $dx\,dy=n\,dS$ and the integral (2) may be written

$$\iint_S F(x, y, z)\, n\, dS\,; \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

but questions of sign arise when n is negative and the relations

between the forms (2) and (3) must be investigated. In some of the most important applications the double integral (2) has to be transformed into a curvilinear integral round C_3, and the relation between the positive directions of integration along C and along C_3 needs elucidation. The following example illustrates the nature of the difficulties (or ambiguities) that occur.

Suppose that the partial derivatives F_y and F_z are continuous when (x, y, z) is any point on S; by Green's Theorem (§ 131 (3))

$$\int_{C_3} F_3(x, y)\,dx = -\iint_{S_3} \frac{\partial F_3(x, y)}{\partial y}\,dx\,dy, \quad \text{...........(4)}$$

provided the curvilinear integral is taken in the positive direction round C_3.

Now
$$\frac{\partial F_3(x, y)}{\partial y} = \frac{\partial F(x, y, z)}{\partial y} + \frac{\partial F(x, y, z)}{\partial z}\frac{\partial z}{\partial y}$$

since x is constant. The value of $\partial z/\partial y$ may be obtained from the equation of the surface S, namely, $z = h(x, y)$ or its equivalent $\varphi(x, y, z) = 0$; thus,

$$\frac{\partial \varphi}{\partial y} + \frac{\partial \varphi}{\partial z}\frac{\partial z}{\partial y} = 0 \text{ or } \frac{\partial z}{\partial y} = -\frac{\varphi_y}{\varphi_z} = -\frac{m}{n},$$

because the direction cosines of the normal are proportional to $\varphi_x, \varphi_y, \varphi_z$ respectively. Hence

$$\int_{C_3} F_3(x, y)\,dx = \iint_{S_3} \left(m\frac{\partial F(x, y, z)}{\partial z} - n\frac{\partial F(x, y, z)}{\partial y}\right)\frac{dx\,dy}{n}. \quad (5)$$

But, by the definition of a curvilinear integral,

$$\int_{C_3} F_3(x, y)\,dx = \int_C F(x, y, z)\,dx,$$

and $dx\,dy$ = projection of $dS = n\,dS$, if n is positive; therefore

$$\int_C F(x, y, z)\,dx = \iint_S \left(m\frac{\partial F}{\partial z} - n\frac{\partial F}{\partial y}\right)dS. \quad \text{...........(6)}$$

The proof of the equation (6) is, however, unsatisfactory; equation (4) assumes that the direction of integration along C_3 is positive, and we have no guarantee whatever that when (x, y, z) moves along C the projection of the point on the xy plane moves along C_3 in the positive direction. Further, if n were negative, the sign of the double integral in (6) would apparently need to be changed. The whole matter therefore must be considered more carefully.

136. Surface; positive and negative Sides. A small area σ, bounded by a curve γ, is taken on a surface given by the equation $\varphi(x, y, z)=0$ and σ', γ' are the projections of σ, γ on a coordinate plane, say the plane of xy.

Take any point P in σ and let NPN' (Fig. 12) be the normal to the surface at P. The "half-lines" PN and PN' are drawn in opposite directions from the surface σ; one of these directions, say that of PN, is chosen as the positive direction of the normal at P, and PN may be called the positive normal, PN' the negative normal. That side of the surface σ which faces the positive direction of the normal at P will be called the positive side (or face) of the surface σ, the other side of σ which faces the direction PN' being the negative side of the surface. If a surface is closed—the surface of a sphere, for example—the area σ on the side chosen as positive may be supposed to spread out till it covers the whole surface and the whole of that side will be positive; it is not possible to pass from the positive to the negative side without penetrating the surface. If the surface is not closed, such as a spherical cap, it has a bounding edge and it is not possible to pass from one side to the other without crossing the edge.

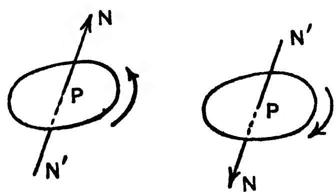

FIG. 12.

The coordinate planes have also two sides. The direction of the positive normal to a coordinate plane is the positive direction of the coordinate axis that is perpendicular to the plane, and the positive side of the plane is that which faces in this direction.

Convention as to sign. The positive direction along the curve γ, that bounds the area σ, is that which is determined by the right-handed screw relation; when the screw advances in the direction PN it twists in the positive direction of rotation round PN (the arrows in the diagram show the relation). This convention agrees with that already adopted for plane curves. Further, the area σ is always assumed to lie on the positive side of the surface and to be positive—that is, measured by a positive number.

Now suppose that σ cannot be cut by a line parallel to the z-axis in more than one point and that n, where n is the cosine of the angle between the positive direction of the z-axis and the positive direction of the normal at any point P in σ, *does not change sign* as P varies in σ so that it is in general either always positive or always negative.

The area σ' which is the projection of σ on the xy plane will always be assumed to lie on the *positive* side of the plane. If Q' is the point on γ' which is the projection of the point Q on γ, it will now be shown that when Q describes γ in the positive direction Q' will describe γ' in the positive or in the negative direction according as n is positive or negative.

A small area σ_1 at any point P on σ projects into an area σ_1' on the xy plane whose measure is (approximately) $n\sigma_1$ and is therefore positive or negative according as n is positive or negative; this relation holds for the complete area σ' since n does not change sign. Now $|\sigma'|$ is given by the double integral

$$\iint dx\,dy, \text{ taken over the area bounded by } \gamma'.$$

When the double integral is transformed into a curvilinear integral round γ' the number that measures the area σ' will be positive or negative according as the direction of integration is positive or negative; in other words, Q' describes γ' in the positive or in the negative direction according as n is positive or negative.

The same considerations apply when the projection is made on the other coordinate planes, the projections of the area σ being always supposed to lie on the positive side of a coordinate plane. The change from one coordinate plane to another is made by the symmetrical change of x, y, z; to pass from the xy to the yz plane, put y for x and z for y, etc.

137. Stokes's Theorem. Consider equations (4), (5) and (6) of § 135. If n is positive no change is needed; when the direction of integration round C is positive, as it is always assumed to be, so is that round C_3. If, however, n is negative, Q' passes round C_3 in the negative direction. But in Green's Theorem the direction round C_3 must be positive, and therefore

in equation (4) the sign of the curvilinear integral must be changed so that the equation becomes

$$-\int_{C_3} F_3(x, y)dx = -\iint_{S_3} \frac{\partial F_3(x, y)}{\partial y} dx\, dy$$

where now the integration round C_3 is in the positive direction. We thus find in this case

$$\int_{C_3} F_3(x, y)dx = \iint_{S_3} \frac{\partial F_3(x, y)}{\partial y} dx\, dy. \quad \text{............(4')}$$

The equation (5) thus becomes

$$\int_{C_3} F_3(x, y)dx = -\iint_{S_3} \left(m\frac{\partial F(x, y, z)}{\partial z} - n\frac{\partial F(x, y, z)}{\partial y}\right)\frac{dx\, dy}{n}. \quad (5')$$

Now, however, $dx\, dy = -n\, dS$ so that we get

$$\int_C F(x, y, z)dx = \iint_S \left(m\frac{\partial F}{\partial z} - n\frac{\partial F}{\partial y}\right) dS \quad \text{............(6)}$$

which is the same equation as before.

Suppose now that $G(x, y, z)$ and $H(x, y, z)$ are single-valued and continuous and have continuous partial derivatives G_z, G_x and H_x, H_y when (x, y, z) is any point on S; it may be proved in the same way, or it may be deduced by symmetry from equation (6), that

$$\int_C G(x, y, z)dy = \iint_S \left(n\frac{\partial G}{\partial x} - l\frac{\partial G}{\partial z}\right) dS \quad \text{............(7)}$$

$$\int_C H(x, y, z)dz = \iint_S \left(l\frac{\partial H}{\partial y} - m\frac{\partial H}{\partial x}\right) dS, \quad \text{.........(8)}$$

and therefore, by addition of (6), (7), (8),

$$\int_C (F dx + G dy + H dz)$$

$$= \iint_S \left\{l\left(\frac{\partial H}{\partial y} - \frac{\partial G}{\partial z}\right) + m\left(\frac{\partial F}{\partial z} - \frac{\partial H}{\partial x}\right) + n\left(\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}\right)\right\} dS. \quad \text{...(9)}$$

Equation (9) is Stokes's Theorem for transforming a curvilinear integral round a curve C into a surface integral over a surface on which C lies.

In equation (9) the surface $\varphi(x, y, z) = 0$ appears only through the numbers l, m, n which are the direction cosines of a normal to $\varphi = 0$; any other surface therefore which satisfied the conditions to which $\varphi = 0$ has been subjected might be taken as that on which the curve C lies.

It has been assumed that the surface is not met by any line parallel to the coordinate axes in more than one point; but this condition is obviously unnecessary if the surface can be divided into a finite number of partial surfaces each of which satisfies the condition just stated. The curvilinear integrals along any curve that is introduced would, as in the case of Green's Theorem, § 131, cancel and the sum of the surface integrals over the partial surfaces would be the integral over the given surface.

Again, it is not necessary that the derivatives φ_x, φ_y, φ_z should be everywhere continuous, provided they are in general continuous. For example, the surface might be a tetrahedron; along the edges the derivatives φ_x, φ_y, φ_z are double-valued, but they are continuous in each plane of the tetrahedron.

138. Green's General Theorem. Let u, v, w be single-valued functions of x, y, z which, with the derivatives $\partial u/\partial x$, $\partial v/\partial y$, $\partial w/\partial z$, are continuous throughout the volume W bounded by a surface S, and let l, m, n be the direction cosines of the normal PP' at a point $P(x, y, z)$ on S, where PP' is the *inward normal*, that is, the half line from P to a point P' near P and inside W. The following relation holds:

$$\iiint_W \left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right) dx\,dy\,dz = -\iint_S (lu+mv+nw)\,dS \quad \ldots(1)$$

where dS is the element of the surface S at (x, y, z).

Let Fig. 13 represent a section of the surface by a plane

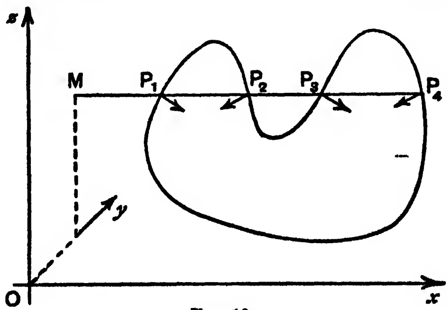

FIG. 13.

parallel to the zx plane. The parallel to the x-axis through point $(0, y, z)$ will meet the surface in 2 or 4 or 6 or, in general, in an even number of points, say P_1, P_2, P_3, P_4.

Now let $MP_1=x_1$, $MP_2=x_2$, $MP_3=x_3$, $MP_4=x_4$; then

$$\iiint \frac{\partial u}{\partial x}\,dx\,dy\,dz=\iint\{-u(x_1, y, z)+u(x_2, y, z)$$
$$-u(x_3, y, z)+u(x_4, y, z)\}\,dy\,dz.$$

A cuboid with base $dy\,dz$ and with lateral surfaces parallel to the x-axis will cut out of the surface S the elements dS_1, dS_2, dS_3, dS_4 at P_1, P_2, P_3, P_4 respectively. If the direction cosines of the inward normals at P_1, P_2, P_3, P_4 are distinguished by the suffixes 1, 2, 3, 4 respectively, then

$$dy\,dz=l_1dS_1=-l_2dS_2=l_3dS_3=-l_4dS_4$$

because $dy\,dz$ is positive and the angle between MP_4 and the normal is acute at P_1 and P_3, obtuse at P_2, P_4. Hence, when expressed in terms of dS, the element of the double integral is of the form $-lu(x, y, z)\,dS$, so that

$$\iiint_w \frac{\partial u}{\partial x}\,dx\,dy\,dz=-\iint_s lu\,dS.$$

The other two triple integrals in (1) may obviously be treated in the same way as the first triple integral, and the equation (1) follows.

Now let $\qquad u=U\dfrac{\partial V}{\partial x},\ v=U\dfrac{\partial V}{\partial y},\ w=U\dfrac{\partial V}{\partial z},$

and let $\qquad \nabla^2V=\dfrac{\partial^2V}{\partial x^2}+\dfrac{\partial^2V}{\partial y^2}+\dfrac{\partial^2V}{\partial z^2};$

then

$$\iiint\left(\frac{\partial U}{\partial x}\frac{\partial V}{\partial x}+\frac{\partial U}{\partial y}\frac{\partial V}{\partial y}+\frac{\partial U}{\partial z}\frac{\partial V}{\partial z}\right)dx\,dy\,dz+\iiint U\nabla^2V\,dx\,dy\,dz$$

$$=-\iint U\left(l\frac{\partial V}{\partial x}+m\frac{\partial V}{\partial y}+n\frac{\partial V}{\partial z}\right)dS \quad\text{.........}(2)$$

$$=-\iint U\frac{\partial V}{\partial \nu}\,dS \quad\text{.........}(3)$$

where (*E.T.* p. 219, (3)) $\partial V/\partial \nu$ is the derivative of V in the direction of the inward normal.

Equations (2) and (3) give Green's Theorem. The form (1) was given independently by Ostrogradsky in a memoir read before the St. Petersburg Academy in the same year (1828) as that in which Green's Essay was published. For a short note on the history of Green's and similar theorems see *Proceedings of the Edinburgh Mathematical Society*, vol. 8, pp. 2-5.

In (3) let $V=U$; then

$$\iiint\left\{\left(\frac{\partial U}{\partial x}\right)^2+\left(\frac{\partial U}{\partial y}\right)^2+\left(\frac{\partial U}{\partial z}\right)^2\right\}dx\,dy\,dz+\iiint U\nabla^2 U\,dx\,dy\,dz$$
$$=-\iint U\frac{\partial U}{\partial \nu}\,dS. \quad \ldots\ldots(4)$$

In (3) interchange U and V and subtract the members of the equation thus obtained from the corresponding members of (3); thus

$$\iiint\left\{U\nabla^2 V-V\nabla^2 U\right\}dx\,dy\,dz=-\iint\left(U\frac{\partial V}{\partial \nu}-V\frac{\partial U}{\partial \nu}\right)dS. \ldots(5)$$

In equations (1) ... (5) the sign (−) before the surface integral will be changed to (+) if the outward normal is chosen instead of the inward.

139. Transformation of $\nabla^2 V$. Suppose that in equation (2) of § 138 the function U is zero on the surface S that bounds the space W; the equation will then become

$$\iiint_W\left[\sum\frac{\partial U}{\partial x}\frac{\partial V}{\partial x}\right]dx\,dy\,dz=-\iiint_W U\nabla^2 V\,dx\,dy\,dz. \ldots\ldots(1)$$

Now let x, y, z be changed to u, v, w where

$$x=f(u,v,w),\ \ y=g(u,v,w),\ \ z=h(u,v,w),$$

and the surfaces $u=$ const., $v=$ const., $w=$ const. are orthogonal.

If ds is the segment joining the points

$$(x, y, z) \text{ and } (x+dx, y+dy, z+dz)$$
$$ds^2=\Sigma dx^2=\sum_{f,g,h}(f_u du+f_v dv+f_w dw)^2,$$

or $$ds^2=\sum_{u,v,w}(f_u^2+g_u^2+h_u^2)du^2=\varrho_1^2 du^2+\varrho_2^2 dv^2+\varrho_3^2 dw^2$$

where $\varrho_1^2=f_u^2+g_u^2+h_u^2,\ \ \varrho_2^2=f_v^2+g_v^2+h_v^2,\ \ \varrho_3^2=f_w^2+g_w^2+h_w^2$
because $f_u f_v+g_u g_v+h_u h_v=0,\ f_v f_w+\ldots=0,\ f_w f_u+\ldots=0$ ...(2)
since the surfaces are orthogonal.

If ds_1, ds_2, ds_3 are the elementary segments of the normals to u, v, w respectively,

$$ds_1=\varrho_1 du,\ \ ds_2=\varrho_2 dv,\ \ ds_3=\varrho_3 dv$$

where ϱ_1, ϱ_2, ϱ_3 are *positive*, and the new element of volume is

$$ds_1 ds_2 ds_3=\varrho_1\varrho_2\varrho_3\,du\,dv\,dw.$$

(Note that $\varrho_1\varrho_2\varrho_3=|J|$, where J is the Jacobian of x, y, z

with respect to u, v, w; this may be seen by forming the square of J.)

From the equations

$$dx = f_u du + f_v dv + f_w dw,\quad dy = g_u du + \ldots,\quad dz = h_u du + \ldots$$

we find

$$\varrho_1^2 du = f_u dx + g_u dy + h_u dz,\quad \varrho_2^2 dv = f_v dx + \ldots,\quad \varrho_3^2 dw = f_w dx + \ldots$$

so that

$$\frac{\partial u}{\partial x} = \frac{f_u}{\varrho_1^2},\quad \frac{\partial u}{\partial y} = \frac{g_u}{\varrho_1^2},\quad \frac{\partial u}{\partial z} = \frac{h_u}{\varrho_1^2},\quad \frac{\partial v}{\partial x} = \frac{f_v}{\varrho_2^2} \ldots,\quad \frac{\partial w}{\partial x} = \frac{f_w}{\varrho_3^2} \ldots.$$

Hence

$$\frac{\partial U}{\partial x} = \frac{\partial U}{\partial u}\frac{f_u}{\varrho_1^2} + \frac{\partial U}{\partial v}\frac{f_v}{\varrho_2^2} + \frac{\partial U}{\partial w}\frac{f_w}{\varrho_3^2},\quad \frac{\partial U}{\partial y} = \ldots,\quad \frac{\partial U}{\partial z} = \ldots$$

so that

$$\sum \frac{\partial U}{\partial x}\frac{\partial V}{\partial x} = \frac{1}{\varrho_1^2}\frac{\partial U}{\partial u}\frac{\partial V}{\partial u} + \frac{1}{\varrho_2^2}\frac{\partial U}{\partial v}\frac{\partial V}{\partial v} + \frac{1}{\varrho_3^2}\frac{\partial U}{\partial w}\frac{\partial V}{\partial w}$$

by using equations (2).

Thus the integral on the left of (1) becomes

$$\iiint_{W'} \left\{ \sum \frac{1}{\varrho_1^2}\frac{\partial U}{\partial u}\frac{\partial V}{\partial u} \right\} \varrho_1 \varrho_2 \varrho_3 \, du\, dv\, dw$$

$$= \iiint_{W'} \left\{ \frac{\varrho_2\varrho_3}{\varrho_1}\frac{\partial U}{\partial u}\frac{\partial V}{\partial u} + \frac{\varrho_3\varrho_1}{\varrho_2}\frac{\partial U}{\partial v}\frac{\partial V}{\partial v} + \frac{\varrho_1\varrho_2}{\varrho_3}\frac{\partial U}{\partial w}\frac{\partial V}{\partial w} \right\} du\, dv\, dw \quad (3)$$

where W' represents the new field of integration.

Now

$$\frac{\varrho_2\varrho_3}{\varrho_1}\frac{\partial U}{\partial u}\frac{\partial V}{\partial u} = \frac{\partial}{\partial u}\left\{ U\left(\frac{\varrho_2\varrho_3}{\varrho_1}\frac{\partial V}{\partial u}\right)\right\} - U\frac{\partial}{\partial u}\left\{\frac{\varrho_2\varrho_3}{\varrho_1}\frac{\partial V}{\partial u}\right\}.$$

If the transformation of § 138 is now applied, u, v, w simply taking the place of x, y, z, the surface integral into which the first term on the right of the equation is transformed will be zero since U is zero on the bounding surface. Hence the integral (3) is equal to

$$-\iiint_{W'} U\left\{\frac{\partial}{\partial u}\left(\frac{\varrho_2\varrho_3}{\varrho_1}\frac{\partial V}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{\varrho_3\varrho_1}{\varrho_2}\frac{\partial V}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{\varrho_1\varrho_2}{\varrho_3}\frac{\partial V}{\partial w}\right)\right\} du\, dv\, dw \quad (4)$$

which is thus the form taken by the integral on the left of (1), U, V being now expressed in terms of u, v, w.

The integral on the right of (1) is transformed into

$$-\iiint_{W'} \left[U\nabla^2 V\right] \varrho_1\varrho_2\varrho_3\, du\, dv\, dw. \quad \ldots\ldots\ldots\ldots\ldots(5)$$

The only restriction to which U is subjected is that it be continuous and vanish on the bounding surface and therefore, since the integrands of the integrals (4) and (5) are continuous, it is necessary that the coefficient of U in each integral be the same. Hence $\nabla^2 V$ becomes

$$\frac{1}{\varrho_1\varrho_2\varrho_3}\left\{\frac{\partial}{\partial u}\left(\frac{\varrho_2\varrho_3}{\varrho_1}\frac{\partial V}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{\varrho_3\varrho_1}{\varrho_2}\frac{\partial V}{\partial v}\right)+\frac{\partial}{\partial w}\left(\frac{\varrho_1\varrho_2}{\varrho_3}\frac{\partial V}{\partial w}\right)\right\}. \quad \ldots(6)$$

The above transformation is due to Jacobi. If $x=f(u, v)$, $y=g(u, v)$ and $z=w$ (so that z is only *formally* changed; compare *E.T.* p. 237, equation (8)); we have

$$f_w=0=g_w,\ h_u=0=h_v,\ h_w=1=\varrho_3,\ \partial^2V/\partial w^2=\partial^2V/\partial z^2$$

and $$\nabla^2 V=\frac{1}{\varrho_1\varrho_2}\left\{\frac{\partial}{\partial u}\left(\frac{\varrho_2}{\varrho_1}\frac{\partial V}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{\varrho_1}{\varrho_2}\frac{\partial V}{\partial v}\right)\right\}+\frac{\partial^2 V}{\partial z^2},$$

where $\partial^2V/\partial z^2$ is zero if V is independent of z.

140. Worked Examples. The following examples give some illustrations of the theorems of Green and Stokes.

Ex. 1. If $H_y=G_z$, $F_z=H_x$, $G_x=F_y$ for every point (x, y, z) in a volume V, bounded by a closed surface S, and if α and β are any two points in V, prove that the integral

$$\int_{\alpha\beta}(F\,dx+G\,dy+H\,dz)$$

has the same value for every path $\alpha\beta$, provided the path lies in V.

As in § 131, Ex. 1, take two paths $\alpha\gamma\beta$ and $\alpha\delta\beta$ lying in V and having no common points except α and β. If C denote the curve $\alpha\gamma\beta\delta\alpha$, the curvilinear integral round C is zero by § 137, (9), and therefore the integrals along $\alpha\gamma\beta$ and $\alpha\delta\beta$ are equal.

Ex. 2. If the curvilinear integral in Ex. 1 is independent of the path $\alpha\beta$ where α and β are any two points in V and the path lies in V, show that $H_y=G_z$, $F_z=H_x$, $G_x=F_y$.

Proceed as in § 131, Ex. 2.

Ex. 3. If $H_y=G_z$, $F_z=H_x$, $G_x=F_y$ for every point (x, y, z) in V, prove that $F\,dx+G\,dy+H\,dz$ is a complete differential.

Let $P(\xi, \eta, \zeta)$ be any point in V. We can choose the points $L(a, b, c)$, $M(\xi, b, c)$, $N(\xi, \eta, c)$ so that the path $LMNP$ (or p) lies in V. Now let (ξ, η, ζ) be defined by the integral

$$f(\xi, \eta, \zeta)=\int_p(F\,dx+G\,dy+H\,dz)+\text{const.}$$

$$=\int_a^\xi F(x, b, c)dx+\int_b^\eta G(\xi, y, c)dy+\int_c^\zeta H(\xi, \eta, z)dz+\text{const.}$$

Differentiate with respect to ξ and note that

$$\frac{\partial G(\xi, y, c)}{\partial \xi}=\frac{\partial F(\xi, y, c)}{\partial y}, \quad \frac{\partial H(\xi, \eta, z)}{\partial \xi}=\frac{\partial F(\xi, \eta, z)}{\partial z}.$$

ence

$$\frac{\partial f}{\partial \xi}=F(\xi, b, c)+[F(\xi, \eta, c)-F(\xi, b, c)]+[F(\xi, \eta, \zeta)-F(\xi, \eta, c)]$$
$$=F(\xi, \eta, \zeta).$$

Similarly $\quad \partial f/\partial \eta=G(\xi, \eta, \zeta), \ \partial f/\partial \zeta=H(\xi, \eta, \zeta)$ so that

$$F(\xi, \eta, \zeta)d\xi+G(\xi, \eta, \zeta)d\eta+H(\xi, \eta, \zeta)d\zeta=df(\xi, \eta, \zeta),$$

and for ξ, η, ζ we may put x, y, z respectively.

Ex. 4. If u, v, w satisfy the conditions of § 138 and if K is a surface inside W bounded by a closed curve C, under what conditions will the surface integral

$$\iint_K (lu+mv+nw)\,dS \quad \text{..............................(i)}$$

depend solely on the curve C and not on the particular surface K on which C lies ?

The numbers l, m, n are the direction cosines of the *negative* normal at a point P on the surface K (§ 136) and the integration is taken over the positive face of K. Now let S be a closed surface formed by two surfaces K_1 and K_2 that lie in W and pass through C but contain no other common points than those on C; denote by V_1 the volume enclosed by S.

If the integral (i) depends only on C then

$$\iint_{K_1} (lu+mv+nw)\,dS=\iint_{K_2} (lu+mv+nw)\,dS$$

and therefore, if K_1 is on the positive face of S,

$$\iint_S (lu+mv+nw)\,dS=0;$$

because in the integral over the part K_2 of S the normal at a point P on K_2 is *an inward normal to* S, that is, is a positive normal to K_2 so that

$$\iint_S (lu+mv+nw)\,dS=\int_{K_1} (lu+mv+nw)\,dS-\int_{K_2} (lu+mv+nw)\,dS.$$

Hence by § 138, equation (1),

$$\iiint_{V_1}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)dx\,dy\,dz=\iint_S (lu+mv+nw)\,dS=0.$$

Now the surface S is arbitrary since K_1 and K_2 are any surfaces, and therefore the triple integral cannot be zero unless its integrand, which is a continuous function, is zero. Hence the condition

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \quad \text{..................................(ii)}$$

must be satisfied at every point in W if the integral (i) is to depend solely on C and not on K; that is, condition (ii) is *necessary*.

That condition (ii) is also sufficient may be shown by finding, if it is satisfied, the line integral round C that is equal to the surface integral (i).

By Stokes's Theorem we have to find F, G, H so that

$$\frac{\partial H}{\partial y}-\frac{\partial G}{\partial z}=u,\quad \frac{\partial F}{\partial z}-\frac{\partial H}{\partial x}=v,\quad \frac{\partial G}{\partial x}-\frac{\partial F}{\partial y}=w. \quad \text{...............(iii)}$$

Now if F_1, G_1, H_1 are values of F, G, H respectively that satisfy (iii) so will the functions

$$F=F_1=\frac{\partial \psi}{\partial x},\quad G=G_1+\frac{\partial \psi}{\partial y},\quad H=H_1+\frac{\partial \psi}{\partial z} \quad \text{...............(iv)}$$

where ψ is a single-valued function of x, y, z which has continuous first and second derivatives; because

$$\frac{\partial H}{\partial y}-\frac{\partial G}{\partial z}=\left(\frac{\partial H_1}{\partial y}+\frac{\partial^2 \psi}{\partial y \partial z}\right)-\left(\frac{\partial G_1}{\partial z}+\frac{\partial^2 \psi}{\partial y \partial z}\right)=u, \text{ etc.}$$

The functions F_1, G_1, H_1 may be taken to be

$$F_1=\int_c^z v\,dz+\varphi_1(x, y),\; G_1=-\int_c^z u\,dz+\varphi_2(x, y),\; H_1=0 \quad \text{.........(v)}$$

where c is a constant and φ_1, φ_2 are functions of x and y. These equations give

$$\frac{\partial H_1}{\partial y}-\frac{\partial G_1}{\partial z}=u,\quad \frac{\partial F_1}{\partial z}-\frac{\partial H_1}{\partial x}=v \text{ and}$$

$$\frac{\partial G_1}{\partial x}-\frac{\partial F_1}{\partial y}=-\int_c^z\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)dz+\frac{\partial \varphi_2}{\partial x}-\frac{\partial \varphi_1}{\partial y}=\int_c^z \frac{\partial w}{\partial z}dz+\frac{\partial \varphi_2}{\partial x}-\frac{\partial \varphi_1}{\partial y}$$

by equation (ii). Hence

$$\frac{\partial G_1}{\partial x}-\frac{\partial F_1}{\partial y}=w(x, y, z)-w(x, y, c)+\frac{\partial \varphi_2}{\partial x}-\frac{\partial \varphi_1}{\partial y}=w(x, y, z)$$

if $$\frac{\partial \varphi_2}{\partial x}-\frac{\partial \varphi_1}{\partial y}=w(x, y, c). \quad \text{..(vi)}$$

Let one of the functions φ_1, φ_2 be chosen arbitrarily; the other is then determined by the equation (vi) and the functions F_1, G_1, H_1 satisfy equations (iii). Hence the functions F, G, H given by (iv) are such that

$$\iint_S (lu+mv+nw)\,dS=\int_C (F\,dx+G\,dy+H\,dz).$$

Solid Angle. Definition. The solid angle subtended at a point A in space by a surface S, bounded by a closed curve C, is measured by the area intercepted on the sphere with centre A and unit radius by the cone with vertex A and the surface S as base; the measure is positive or negative according as the positive or negative face of the surface is seen from A.

If the surface S is closed (like an ellipsoid), the solid angle will be measured by the complete surface of the unit sphere when A is inside S but will be zero when A is outside S. These two results may be obtained from the expression (i) of Ex. 5

as limiting cases ; the definition assumes that S is not a closed surface but is bounded by a closed curve.

Ex. 5. Expression for a solid angle.

Let A be (a, b, c), P any point (x, y, z) in the small area δS and θ the angle between the *positive normal* PN to δS and the direction PA *from* P *to* A. Let $\delta\omega$ be the area intercepted on the sphere with centre A and unit radius by the cone with vertex A and base δS ; the area of the section of the cone by the plane through P perpendicular to PA is (approximately) $\delta S\,|\cos\theta|$ in numerical value where δS is *positive*. But, if $|PA|=r$, we have, by geometry, $r^2\,|\delta\omega|=\delta S\,|\cos\theta|$ and therefore, in sign and magnitude, $r^2\,\delta\omega=\delta S\cos\theta$ so that

$$\omega=\iint_S \frac{\cos\theta}{r^2}\,dS. \qquad \text{.................................(i)}$$

If l, m, n are the direction cosines of PN, we have

$$\cos\theta=\{l(a-x)+m(b-y)+n(c-z)\}/r$$

since the direction cosines of PA are $(a-x)/r$, $(b-y)/r$, $(c-z)/r$, so that

$$\omega=\iint_S\left\{\frac{l(a-x)+m(b-y)+n(c-z)}{r^3}\right\}dS. \qquad \text{(ii)}$$

If $u=(a-x)/r^3$, $v=(b-y)/r^3$, $w=(c-z)/r^3$, it is easy to verify that

$$\frac{\partial u}{\partial x}+\frac{\partial w}{\partial y}+\frac{\partial w}{\partial z}=0,$$

so that (Ex. 4) ω is independent of the surface S and depends only on the bounding curve.

EXERCISES XVI.

1. Change the order of integration in the integral

$$\int_0^{\frac{ab}{c}}dy\int_0^{\frac{a}{b}\sqrt{(b^2-y^2)}}F(x,y)\,dx.$$

2. When the field of integration is the triangle given by $y=0$, $y=x$ and $x=1$, show that

$$\iint\sqrt{(4x^2-y^2)}\,dx\,dy=\tfrac{1}{3}\left(\frac{\pi}{3}+\frac{\sqrt{3}}{2}\right).$$

3. $$\int_0^{-}dx\int_0^{b}\frac{dy}{(x^2+y^2+c^2)^{\frac{3}{2}}}=\frac{1}{c}\tan^{-1}\left\{\frac{ab}{c\sqrt{(a^2+b^2+c^2)}}\right\}.$$

4. When the field of integration is the circle $x^2+y^2=2ay$, show that

$$\iint\sqrt{(4ay-x^2)}\,dx\,dy=\tfrac{1}{3}(3\pi+8)a^3.$$

5. The integral of $(x\sin\alpha-y\cos\alpha)^2$, taken over the ellipse $x^2/a^2+y^2/b^2=1$, is equal to

$$\tfrac{1}{4}\pi ab(a^2\sin^2\alpha+b^2\cos^2\alpha).$$

If the integral is taken over the rectangle given by $x=a$, $x=-a$ and $y=b$, $y=-b$, its value is

$$\tfrac{4}{3}ab(a^2\sin^2\alpha+b^2\cos^2\alpha).$$

6. If $I = \iint \frac{dx\,dy}{\sqrt{(x^2+y^2+c^2)}}$, taken over the circle $x^2+y^2=a^2$, find I and show that

$$-\frac{\partial I}{\partial c} = 2\pi\left\{1 - \frac{c}{\sqrt{(a^2+c^2)}}\right\}.$$

7. $\iiint (ax^2+by^2+cz^2)\,dx\,dy\,dz$, taken through the sphere $x^2+y^2+z^2=R^2$ is $\frac{4}{15}\pi(a+b+c)R^5$.

8. $\iiint_V z\,dx\,dy\,dz = \frac{\pi}{4}c^4 \cot\alpha \cot\beta$

where V is the volume bounded by the cone $z^2 = x^2\tan^2\alpha + y^2\tan^2\beta$ and the planes $z=0$, $z=c$.

9. If V' is that part of the volume V of Example 8 for which x, y, and z are positive, show that

$$xyz\,dx\,dy\,dz = \tfrac{1}{48}\, c^6 \cot^2\alpha \cot^2\beta.$$

10. If p is the perpendicular from (x, y, z) on the diameter of the ellipsoid $x^2/a^2+y^2/b^2+z^2/c^2=1$ which is inclined to the axes at the angles α, β, γ, prove that

$$\iiint p^2\,dx\,dy\,dz = \frac{4\pi}{15}abc(a^2\sin^2\alpha + b^2\sin^2\beta + c^2\sin^2\gamma)$$

where the integration is taken through the ellipsoid.

11. The value of $\iiint z^2\,dx\,dy\,dz$, taken through the volume common to the sphere $x^2+y^2+z^2=a^2$ and the cylinder $x^2+y^2=ax$, is

$$\frac{4}{15}\left(\frac{\pi}{2} - \frac{8}{15}\right)a^5.$$

12. The mean value of $(ax+by+cz)^{2n}$, where n is a positive integer, over the surface of the sphere $x^2+y^2+z^2=1$ is

$$(a^2+b^2+c^2)^n/(2n+1)$$

and the mean value over the volume of the sphere is

$$3(a^2+b^2+c^2)^n/(2n+1)(2n+3).$$

13.* $\int_0^1 y^{m-1}\,dy \int_0^{1-y} x^{2l-1}(1+x)^{m+n-1}(1-x-y)^{n-1}\,dx$

$$= \tfrac{1}{2}\Gamma(l)\Gamma(m)\Gamma(n)/\Gamma(l+m+n),\ l \geqq 1,\ m \geqq 1,\ n \geqq 1.$$

14. If m, n, p are each not less than unity,

$$\int_0^a (a-x)^{m-1}\,dx \int_0^x (x-y)^{n-1}\,dy \int_0^y (y-z)^{p-1} f(z)\,dz$$
$$= \frac{\Gamma(m)\Gamma(n)\Gamma(p)}{\Gamma(m+n+p)} \int_0^a (a-z)^{m+n+p-1} f(z)\,dz.$$

* l, m, n are taken to be each not less than 1, so that the integrand may be a bounded function of x and y. When the improper integral has been defined it will be seen that the result holds if l, m, n are each positive. A similar remark is applicable in the case of other examples.

15. If $m \geqq 0$, prove that the integral

$$\iint\left(1-\frac{x^2}{a^2}-\frac{y^2}{b^2}\right)^m f(Ax+By)\,dx\,dy,$$

taken over the ellipse $x^2/a^2+y^2/b^2=1$, is equal to

$$B(\tfrac{1}{2}, m+1)ab\int_{-1}^{1}(1-x^2)^{m+\frac{1}{2}}f(kx)dx,\quad k=(A^2a^2+B^2b^2)^{\frac{1}{2}}.$$

16. $$\int_0^{\frac{\pi}{2}}d\varphi\int_0^{\frac{\pi}{2}}F(1-\sin\theta\cos\varphi)\sin\theta\,d\theta=\frac{\pi}{2}\int_0^1 F(x)\,dx.$$

17. $$\int_0^{2\pi}d\varphi\int_0^{\pi}\left(\frac{\sin^2\theta\cos^2\varphi}{a^2}+\frac{\sin^2\theta\sin^2\varphi}{b^2}+\frac{\cos^2\theta}{c^2}\right)^{-\frac{3}{2}}\sin\theta\,d\theta=4\pi abc.$$

18. (i) If $c>a$, show that the integral

$$\iint\frac{(c-x)\,dx\,dy}{(c-x)^2+y^2},$$

taken over the ellipse $x^2/a^2+y^2/b^2=1$, is equal to

$$\frac{2\pi ab}{a^2-b^2}\left\{c-\sqrt{(c^2-a^2+b^2)}\right\}.$$

(ii) If $a>h>0$, show, that the integral

$$\iiint\frac{(a-z)\,dx\,dy\,dz}{\{x^2+y^2+(a-z)^2\}^{\frac{3}{2}}}$$

taken throughout the volume bounded by the cylinder $x^2+y^2=c^2$ and the planes $z=h$, $z=-h$, is equal to

$$\pi\{[c^2+(a-h)^2]^{\frac{1}{2}}-[c^2+(a+h)^2]^{\frac{1}{2}}+2h\}.$$

19. Prove that when the double integral in x and y is taken over the positive quadrant of the circle $x^2+y^2=1$,

$$\text{(i)}\quad\iint\frac{dx\,dy}{\sqrt{(1-x^2-y^2)}}=\frac{\pi}{2},$$

and deduce that, if $x=\sin\theta\sqrt{(1-m^2\sin^2\varphi)}$, $y=\sin\varphi\sqrt{(1-n^2\sin^2\theta)}$, where $m^2+n^2=1$ and $0\leqq\theta\leqq\pi/2$, $0\leqq\varphi\leqq\pi/2$,

$$\text{(ii)}\quad\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\frac{(m^2\cos^2\varphi+n^2\cos^2\theta)d\theta\,d\varphi}{\sqrt{(1-m^2\sin^2\varphi)}\sqrt{(1-n^2\sin^2\theta)}}=\frac{\pi}{2}.$$

Give a geometrical interpretation of the integral (i).

[The integral (i) is not a proper integral. See, however, p. 380.]

20. If $$F(m)=\int_0^{\frac{\pi}{2}}\frac{d\theta}{\sqrt{(1-m^2\sin^2\theta)}},\quad E(m)=\int_0^{\frac{\pi}{2}}\sqrt{(1-m^2\sin^2\theta)}d\theta,$$

and $m^2+n^2=1$, deduce, by the help of Example 19, that

$$F(m)E(n)+F(n)E(m)-F(m)F(n)=\frac{\pi}{2}.$$

21. If $x=r\sin\theta\sqrt{(1-m^2\sin^2\varphi)}$, $y=r\sin\varphi\sqrt{(1-n^2\sin^2\theta)}$, $z=r\cos\theta\cos\varphi$ where, as in Example 19, $m^2+n^2=1$, prove that (x, y, z) is a point on a sphere of radius r and change the variables in the integral $\iiint dx\,dy\,dz$, taken over that octant of the sphere $x^2+y^2+z^2=1$ for which x, y, z are all positive, to the variables r, θ, φ.

Deduce the value of the integral in Example 19, (ii).

22. Prove that

$$\int_0^{2\pi} d\varphi \int_0^{\pi} F(a\sin\theta\cos\varphi+b\sin\theta\sin\varphi+c\cos\theta)\sin\theta\,d\theta$$
$$=2\pi\int_{-1}^{1} F(kx)\,dx,\quad k=(a^2+b^2+c^2)^{\frac{1}{2}}.$$

By differentiating with respect to a, b, c other integrals may be derived; thus, differentiating as to c and putting $f(x)$ for $F'(x)$ we find

$$\int_0^{2\pi} d\varphi \int_0^{\pi} f(a\sin\theta\cos\varphi+b\sin\theta\sin\varphi+c\cos\theta)\cos\theta\sin\theta\,d\theta$$
$$=\frac{2\pi c}{k}\int_{-1}^{1} f(kx)\,x\,dx.$$

23. From the sphere $x^2+y^2+z^2=a^2$ a segment is cut off by the plane $z=h>0$; P is any point on the curved surface S, Q any point in the volume V of the smaller segment and C the point $(0, 0, c)$, where $c>a$. Prove

(i) $$U_1=\iint_S \frac{dS}{|CP|}=\frac{2\pi a}{c}\{(c^2-2ch+a^2)^{\frac{1}{2}}-c+a\};$$

(ii) $$U_2=\iiint_V \frac{dx\,dy\,dz}{|CQ|}$$
$$=\frac{2\pi}{3c}\{(c^2-2ch+a^2)^{\frac{3}{2}}+a^3\}-\frac{\pi}{3}(2c^2-6ch+3h^2+3a^2).$$

If the segment is made by the plane $z=-k$, where $k>0$, show that the values of U_1 and U_2 are obtained by writing $-k$ for h.

24. If $F(x, y, z)=x^{l-1}y^{m-1}z^{n-1}f\left\{\left(\frac{x}{a}\right)^p+\left(\frac{y}{b}\right)^q+\left(\frac{z}{c}\right)^r\right\}$, where a, b, c are positive and the indices are such that $F(x, y, z)$ is continuous, show that the integral of $F(x, y, z)$, taken through the part of the volume bounded by the surface

$$(x/a)^p+(y/b)^q+(z/c)^r=1$$

in which x, y, z are all positive, is equal to

$$\frac{a^l b^m c^n}{pqr}\int_0^1 x^{\frac{l}{p}-1}\,dx\int_0^{1-x} y^{\frac{m}{q}-1}\,dy\int_0^{1-x-y} z^{\frac{n}{r}-1} f(x+y+z)\,dz.$$

The example may easily be extended to the case of n variables. The integral may be transformed to one with constant limits (§ 133, Examples 3, 7). If $f(u)=(1-u)^{\alpha}$, the integral can be expressed in terms of Gamma Functions.

25. *Elliptic Coordinates.** The equation $f(t)=0$, where

$$f(t)=1-\frac{x^2}{t}-\frac{y^2}{t-b^2}-\frac{z^2}{t-c^2},\quad b^2<c^2,$$

is satisfied by three real values λ^2, μ^2, ν^2, where

$$\lambda^2>c^2>\mu^2>b^2>\nu^2$$

and the three surfaces $\lambda=$const., $\mu=$const., $\nu=$const. are confocal conicoids which pass through the point (x, y, z) and intersect orthogonally. The values of x, y, z are given by

$$x^2=\frac{\lambda^2\mu^2\nu^2}{b^2c^2},\quad y^2=\frac{(\lambda^2-b^2)(\mu^2-b^2)(\nu^2-b^2)}{b^2(b^2-c^2)},\quad z^2=\frac{(\lambda^2-c^2)(\mu^2-c^2)(\nu^2-c^2)}{c^2(c^2-b^2)}.$$

When λ, μ, ν are fixed, the point (x, y, z) will be uniquely determined if, for example, ν, $\sqrt{(\mu^2-b^2)}$ and $\sqrt{(\lambda^2-c^2)}$ are allowed to take either positive or negative values while λ, μ and the other square roots are kept positive. The numbers λ, μ, ν are called the *elliptic coordinates* of the point (x, y, z).

(i) If $J=\dfrac{\partial(x,y,z)}{\partial(\lambda,\mu,\nu)}$, prove that

$$|J|=\frac{(\lambda^2-\mu^2)(\lambda^2-\nu^2)(\mu^2-\nu^2)}{\sqrt{\{(\lambda^2-b^2)(\lambda^2-c^2)(\mu^2-b^2)(c^2-\mu^2)(b^2-\nu^2)(c^2-\nu^2)\}}}.$$

(ii) If ds is the distance between the points

$$(\lambda, \mu, \nu) \text{ and } (\lambda+d\lambda, \mu+d\mu, \nu+d\nu),$$

show that

$$\begin{aligned} ds^2 &= dx^2+dy^2+dz^2 \\ &= \frac{(\lambda^2-\mu^2)(\lambda^2-\nu^2)}{(\lambda^2-b^2)(\lambda^2-c^2)}d\lambda^2+\frac{(\lambda^2-\mu^2)(\mu^2-\nu^2)}{(\mu^2-b^2)(c^2-\mu^2)}d\mu^2+\frac{(\lambda^2-\nu^2)(\mu^2-\nu^2)}{(b^2-\nu^2)(c^2-\nu^2)}d\nu^2. \end{aligned}$$

(iii) If ds_1, ds_2 and ds_3 are the respective values of ds when λ, μ and ν alone vary, prove that

$$ds_1=A\,d\lambda,\quad ds_2=B\,d\mu,\quad ds_3=C\,d\nu,$$

where
$$A=\left|\left\{\frac{(\lambda^2-\mu^2)(\lambda^2-\nu^2)}{(\lambda^2-b^2)(\lambda^2-c^2)}\right\}^{\frac12}\right|,\quad B=\left|\left\{\frac{(\lambda^2-\mu^2)(\mu^2-\nu^2)}{(\mu^2-b^2)(c^2-\mu^2)}\right\}^{\frac12}\right|,$$
$$C=\left|\left\{\frac{(\lambda^2-\nu^2)(\mu^2-\nu^2)}{(b^2-\nu^2)(c^2-\nu^2)}\right\}^{\frac12}\right|.$$

Deduce from (iii) the value of $|J|$.

26. If p is the length of the perpendicular from the centre of the ellipsoid $\lambda=$const. on the tangent plane at (λ, μ, ν), prove that

$$p^2=\lambda^2(\lambda^2-b^2)(\lambda^2-c^2)/(\lambda^2-\mu^2)(\lambda^2-\nu^2)$$

and deduce, by expressing the volume of the positive octant of the ellipsoid as the integral $\frac13\iint p\,ds_2\,ds_3$ over the surface, that

$$\int_b^c d\mu\int_0^b\frac{(\mu^2-\nu^2)\,d\nu}{\sqrt{\{(\mu^2-b^2)(c^2-\mu^2)(b^2-\nu^2)(c^2-\nu^2)\}}}=\frac{\pi}{2}.$$

* See Bell's *Coordinate Geometry of Three Dimensions*, Chapter X, for the properties of Confocal Conicoids.

27. Prove the following results:

(i) $\int_c^{\lambda_1} d\lambda \int_b^c d\mu \int_0^b |J|\,d\nu = \frac{\pi}{6}\lambda_1\sqrt{\{(\lambda_1^2-b^2)(\lambda_1^2-c^2)\}}$;

(ii) $\int_b^c d\mu \int_0^b (\mu^2-\nu^2)\sqrt{\left\{\frac{(c^2-\mu^2)(c^2-\nu^2)}{(\mu^2-b^2)(b^2-\nu^2)}\right\}}\,d\nu = \frac{\pi}{6}c^2(c^2-b^2)$;

(iii) $\int_b^c d\mu \int_0^b \frac{(\lambda^2-\mu^2)(\lambda^2-\nu^2)(\mu^2-\nu^2)\,d\nu}{\sqrt{\{(\mu^2-b^2)(c^2-\mu^2)(b^2-\nu^2)(c^2-\nu^2)\}}} = \frac{\pi}{6}\{3\lambda^4-2(b^2+c^2)\lambda^2+b^2c^2\}$.

28. Express $\nabla^2 V$ in terms of the elliptic coordinates λ, μ, ν. If

$$\xi = a^2\int_c^\lambda \frac{d\lambda}{\sqrt{\{(\lambda^2-b^2)(\lambda^2-c^2)\}}}, \quad \eta = a^2\int_b^\mu \frac{d\mu}{\sqrt{\{(\mu^2-b^2)(c^2-\mu^2)\}}},$$

$$\zeta = a^2\int_0^\nu \frac{d\nu}{\sqrt{\{(b^2-\nu^2)(c^2-\nu^2)\}}},$$

show that

$$\nabla^2 V = K\left\{(\mu^2-\nu^2)\frac{\partial^2 V}{\partial\xi^2} + (\lambda^2-\nu^2)\frac{\partial^2 V}{\partial\eta^2} + (\lambda^2-\mu^2)\frac{\partial^2 V}{\partial\zeta^2}\right\}$$

where $\qquad K = a^4/(\lambda^2-\mu^2)(\lambda^2-\nu^2)(\mu^2-\nu^2).$

29. E is the ellipse $x^2/a^2+y^2/b^2=1$ and the equations

$$x^2/\cosh^2 u + y^2/\sinh^2 u = c^2, \quad x^2/\cos^2 v - y^2/\sin^2 v = c^2$$

give respectively ellipses and hyperbolas confocal with E.

If p is the perpendicular from $(0, 0)$ to the tangent at (x, y) on the ellipse u, show that x and y may be expressed in the forms

$$x = c\cosh u\cos v, \quad y = c\sinh u\sin v$$

and then prove

(i) $J = \frac{\partial(x, y)}{\partial(u, v)} = c^2(\cosh^2 u\sin^2 v + \sinh^2 u\cos^2 v)$, $p^2 J = c^4\cosh^2 u\sinh^2 u$;

(ii) $\iint_E p^2\,dx\,dy = \frac{\pi}{4}\left\{ab(a^2+b^2) - (a^2-b^2)^2\cosh^{-1}\frac{a}{\sqrt{(a^2-b^2)}}\right\}$.

30. If $x=\varrho\cos\varphi$, $y=\varrho\sin\varphi$, $z=z$ and if $U=\varrho^{\frac{1}{2}}V$, prove that

$$\varrho^{\frac{1}{2}}\nabla^2 V = \frac{\partial^2 U}{\partial\varrho^2} + \frac{1}{\varrho^2}\frac{\partial^2 U}{\partial\varphi^2} + \frac{\partial^2 U}{\partial z^2} + \frac{U}{4\varrho^2},$$

and if $\qquad \varrho = \frac{b\sinh\lambda}{\cosh\lambda+\cos\omega}, \quad z = \frac{b\sin\omega}{\cosh\lambda+\cos\omega}$

$$\varrho^{\frac{5}{2}}\nabla^2 V = (\sinh\lambda)^2\left(\frac{\partial^2 U}{\partial\lambda^2} + \frac{\partial^2 U}{\partial\omega^2}\right) + \frac{\partial^2 U}{\partial\varphi^2} + \frac{1}{4}U.$$

31. If $x_1 = r\cos\theta_1$, $x_2 = r\sin\theta_1\cos\theta_2$, $x_3 = r\sin\theta_1\sin\theta_2\cos\theta_3$ and $x_4 = r\sin\theta_1\sin\theta_2\sin\theta_3$, show that

$$\frac{\partial(x_1, x_2, x_3, x_4)}{\partial(r, \theta_1, \theta_2, \theta_3)} = r^3\sin^2\theta_1\sin\theta_2,$$

and extend to the case of two sets of n variables $x_1, x_2, \ldots, x_n$ and $r, \theta_1, \ldots, \theta_{n-1}$. Prove that

(i) $x_1^2+x_2^2+\ldots+x_n^2 = r^2$;

(ii) $\frac{\partial(x_1, x_2, x_3, \ldots, x_n)}{\partial(r, \theta_1, \theta_2, \ldots, \theta_{n-1})} = r^{n-1}(\sin\theta_1)^{n-2}(\sin\theta_2)^{n-3}\ldots(\sin\theta_{n-2})$.

32. If

$$x_1 = r\sin\theta\cos\varphi,\ x_2 = r\sin\theta\sin\varphi,\ x_3 = r\cos\theta\cos\psi$$

and $\quad x_4 = r\cos\theta\sin\psi$ then $\Sigma x^2 = r^2$

and

$$\frac{\partial(x_1, x_2, x_3, x_4)}{\partial(r, \theta, \varphi, \psi)} = r^3\cos\theta\sin\theta.$$

33. If the rectangular axes of x, y, z are changed to another set of rectangular axes ξ, η, ζ with a new origin (a_1, a_2, a_3), show that

$$\iint\sqrt{\left\{\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2+1\right\}}\,dx\,dy=\iint\sqrt{\left\{\left(\frac{\partial \zeta}{\partial \xi}\right)^2+\left(\frac{\partial \zeta}{\partial \eta}\right)^2+1\right\}}\,d\xi\,d\eta.$$

The formulae of transformation are

$$x, y, z = a_r + l_r\xi + m_r\eta + n_r\zeta, \quad r = 1, 2, 3.$$

If $p = \partial z/\partial x$, $q = \partial z/\partial y$ and $P = \partial\zeta/\partial\xi$, $Q = \partial\zeta/\partial\eta$, it is not hard to prove that

$$\text{(i)}\ \sqrt{(p^2+q^2+1)} = |\,n_1p + n_2q - n_3\,| \,.\, \sqrt{(P^2+Q^2+1)};$$

$$\text{(ii)}\ \left|\frac{\partial(x, y)}{\partial(\xi, \eta)}\right| . \left|\,n_1p + n_2q - n_3\,\right| = 1.$$

The measure S of a surface is therefore independent of any particular set of coordinate axes.

34. If $x = f(u, v, w)$, $y = g(u, v, w)$, $z = h(u, v, w)$ and if P, A, B, C are the points determined by the parameters

$$(u, v, w),\ (u + du, v, w),\ (u, v + dv, w),\ (u, v, w + dw)$$

respectively, prove that the volume of the tetrahedron $PABC$ is $\frac{1}{6}|\,J\,|\,du\,dv\,dw$ where J is the Jacobian of f, g, h with respect to u, v, w and du, dv, dw are positive.

The volume of the parallelepiped of which PA, PB, PC are conterminous edges is $|\,J\,|\,du\,dv\,dw$. Deduce the transformation of Problem II, § 134. (Compare Exercises VI, 14.)

35. The value of $\iint(lx^2 + my^2 + nz^2)dS$, taken over the surface of the sphere

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = R^2$$

is $\frac{8}{3}\pi(a+b+c)R^3$, the direction cosines l, m, n being those of the *outward* normal to the sphere.

Verify the result by transforming the integral into a triple integral, taken through the sphere.

36. C is the curve given by the equations

$$x^2 + y^2 + z^2 - 2ax - 2ay = 0,\ x + y = 2a;$$

prove that

$$\int_C (y\,dx + z\,dy + x\,dz) = -2\sqrt{2}\,.\,\pi a^2,$$

the path beginning at the point $(2a, 0, 0)$ and lying at first below the plane of xy. Transform the curvilinear integral into a surface integral over the plane area enclosed by C.

37. Verify Example 19, Exercises XIV, by applying Stokes's Theorem to transform the curvilinear integral into a surface integral over the relative portion of the surface of the sphere.

38. On the perpendicular from the centre O of the circle $x^2+y^2=a^2$ to the plane of the circle a point P is taken; if $OP=z>0$, show that the solid angle ω subtended at P by the circle is

$$2\pi\left\{1-\frac{z}{\sqrt{(a^2+z^2)}}\right\}.$$

If $z<0$, what is the value of ω?

39. Prove that, if the symbols have the same meaning as in Example 5, § 140, the integral (i) that measures ω is, when S is a closed surface, 4π or 0 according as A is inside or outside the surface, the *inward* normal being considered the positive normal.

If A is outside S, a line through A will meet S at an even number of points, P_1, P_2, P_3, P_4 say (compare Fig. 13); a cone with vertex A and small vertical angle, having $AP_1 \ldots P_4$ as axis, will intercept areas $\delta S_1, \ldots, \delta S_4$ at $P_1, \ldots, P_4$. If these areas be projected on the unit sphere with centre A, the area intercepted on the sphere by the cone being $\delta\omega$, then

$$\begin{aligned}\delta\omega = \delta S_1 \cos\theta_1/AP_1^2 &= -\,\delta S_2\cos\theta_2/AP_2^2\\ &= \delta S_3\cos\theta_3/AP_3^2 = -\,\delta S_4\cos\theta_4/AP_4^2,\end{aligned}$$

so that the sum of the four elements $\delta S\cos\theta/AP^2$ is zero. For the whole surface it follows that the sum is zero so that ω is zero.

If A is inside it is plain that the sum is simply the area of the unit sphere, that is, 4π.

40. If as in § 140, Example 5,

$$\omega=\iint_S \frac{l(a-x)+m(b-y)+n(c-z)}{r^3}\,dS,$$

show that

$$\frac{\partial\omega}{\partial a}=\iint_S\left\{l\frac{\partial}{\partial a}\left(\frac{a-x}{r^3}\right)+m\frac{\partial}{\partial a}\left(\frac{b-y}{r^3}\right)+n\frac{\partial}{\partial a}\left(\frac{c-z}{r^3}\right)\right\}dS.$$

If $F=0$, $G=\dfrac{z-c}{r^3}$, $H=-\dfrac{y-b}{r^3}$, show that

$$\frac{\partial\omega}{\partial a}=\int_C (G\,dy+H\,dz)=\int_C\frac{(z-c)dy-(y-b)dz}{r^3}$$

where C is the curve that bounds S.

Prove in the same way that

$$\frac{\partial\omega}{\partial b}=\int_C\frac{(x-a)dz-(z-c)dx}{r^3},$$

$$\frac{\partial\omega}{\partial c}=\int_C\frac{(y-b)dx-(x-a)dy}{r^3}.$$

CHAPTER XII

IMPROPER INTEGRALS

141. Improper Integrals. The definition of an integral in Chapter IX expressly assumes that the integrand is bounded and the range of integration finite. It is possible, however, to extend the definition so that the integral will still have a value when the integrand is not bounded or the range is not finite; the integral, as thus extended, is called an **Improper** (or **Generalised** or **Infinite**) **Integral** while, for the sake of distinction, the integral of Chapter IX is called a **Proper** (or **Finite**) **Integral**.*

The following preliminary definition is given as it simplifies the expression of conditions in many cases.

Singular Point. A point c in an interval (a, b) is called a singular point of the interval for a function $F(x)$ if $F(x)$ is not bounded in the interval $(c-\delta, c+\delta')$ where δ and δ' are arbitrarily small positive numbers; $\delta=0$ when $c=a$ and $\delta'=0$ when $c=b$. It is often convenient to say that $|F(x)|=\infty$ for $x=c$, but this expression means simply that c is a singular point.

It will be assumed throughout that the number of singular points in any interval is *finite* and therefore, when the range of integration is infinite, that all the singular points can be included in a finite interval. This restriction on the number is

* The term "infinite integral" is in some respects more suggestive than "improper integral," especially because of analogies with "infinite series"; but it seems to be too great a strain on language to describe an integral as infinite when the infinity attaches not to the range but to the integrand. None of the terms is really satisfactory, but that of "improper integral" is in such general use that it seems best to retain it.

not necessary for the existence of the improper integral, but the consideration of an infinite number of singular points is beyond our limits.

Note. The sketch of the improper integral in the *Elementary Treatise* (Chapter XXI) is based on the supposition that the integrand is in general continuous, but the method is equally applicable when the integrand satisfies the conditions for a proper integral. The improper integral of a function over any range will be defined as the limit of a proper integral over a part of that range; *a necessary condition for the existence of an improper integral is therefore that both the corresponding proper integral and its limit should exist.* It will save much tedious repetition to assume once for all that the proper integral and the limit both exist, and this assumption—which should be steadily kept in mind—will be adopted; explicit reference will be made to the proper integral and to the limit only when there seems to be special reason for it.

142. Definitions. A set of definitions will now be given. Take first the case of a finite range (a, b), where $b > a$, and let δ, δ' be two arbitrarily small positive numbers.

Range Finite. If a is the only singular point in (a, b) the improper integral of $F(x)$ over (a, b) is defined by the equation

$$\int_a^b F(x)\,dx = \underset{\delta \to 0}{\mathcal{L}} \int_{a+\delta}^b F(x)\,dx, \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

while if b is the only singular point in (a, b)

$$\int_a^b F(x)\,dx = \underset{\delta \to 0}{\mathcal{L}} \int_a^{b-\delta} F(x)\,dx. \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

The integral is often said "to converge at a" (or at b); again, such an expression as "the integral over (a, b) is convergent" is often used as equivalent to the statement that the improper integral over (a, b) exists. Similar language, borrowed from the theory of infinite series, is used throughout and will require no further explanation.

If c, where $a<c<b$, is the only singular point in (a, b) the improper integral over (a, b) is defined by the equation

$$\int_a^b F(x)\,dx = \underset{\delta \to 0}{\mathcal{L}} \int_a^{c-\delta} F(x)\,dx + \underset{\delta' \to 0}{\mathcal{L}} \int_{c+\delta'}^b F(x)\,dx, \quad \ldots\ldots(3)$$

provided *each* of the limits exists ; in other words, the limit must not depend on any relation between δ and δ'.

This proviso is important, as the following simple case shows:

$$\int_a^{c-\delta} \frac{dx}{x-c} = \log\frac{\delta}{c-a}, \quad \int_{c+\delta'}^{b} \frac{dx}{x-c} = \log\frac{b-c}{\delta'}.$$

Neither of the limits for $\delta \to 0$ or $\delta' \to 0$ exists, so that the improper integral of $1/(x-c)$ over (a, b) does not exist. If, however, we suppose $\delta' = \delta$ the limit for $\delta \to 0$ is the definite number $\log[(b-c)/(c-a)]$, and this limit was called by Cauchy the Principal Value of the integral, in accordance with the definition :

Principal Value. If the integral of $F(x)$ as defined by equation (3) has a definite value when $\delta' = \delta$, but not when δ and δ' tend independently to zero, that value is called the Principal Value of the integral.

This so-called Principal Value is clearly of a very special kind, and we shall make little or no use of it. (For notation, see below.)

The definitions (1), (2), (3) may be supplemented by the following.

If a and b (but no other point in (a, b)) are singular points, take any point c such that $a<c<b$; the integral of $F(x)$ over (a, b) is defined by the equation

$$\int_a^b F(x)\,dx = \mathop{\mathcal{L}}_{\delta\to 0}\int_{a+\delta}^{c} F(x)\,dx + \mathop{\mathcal{L}}_{\delta'\to 0}\int_{c}^{b-\delta'} F(x)\,dx, \quad \text{......(4)}$$

provided (as in (3)) that *each* limit exists.

If there are m singular points $c_1, c_2, \dots, c_m$ in (a, b) where $a \leqq c_1$, $c_1 < c_2, \dots, c_m \leqq b$, and if each of the integrals

$$\int_a^{c_1} F(x)\,dx, \int_{c_1}^{c_2} F(x)\,dx, \dots \int_{c_{m-1}}^{c_m} F(x)\,dx, \int_{c_m}^{b} F(x)\,dx$$

exists in the sense of equation (4), then

$$\int_a^b F(x)\,dx = \int_a^{c_1} F(x)\,dx + \sum_{r=1}^{m-1}\int_{c_r}^{c_{r+1}} F(x)\,dx + \int_{c_m}^{b} F(x)\,dx \quad \text{...(5)}$$

where the first integral on the right disappears if $c_1 = a$ and the last if $c_m = b$.

The case of an infinite range of integration will now be considered.

Range Infinite. Suppose $\xi > a$. The definition is now

$$\int_a^\infty F(x)\,dx = \mathop{\mathcal{L}}_{\xi\to\infty}\int_a^\xi F(x)\,dx \quad\text{.....................(6)}$$

while if ξ' is positive and $-\xi' < a$

$$\int_{-\infty}^a F(x)\,dx = \mathop{\mathcal{L}}_{\xi'\to\infty}\int_{-\xi'}^a F(x)\,dx. \quad\text{..................(7)}$$

If each of the limits in equations (6) and (7) exists when ξ and ξ' tend independently to infinity, then, by definition,

$$\int_{-\infty}^\infty F(x)\,dx = \int_{-\infty}^a F(x)\,dx + \int_a^\infty F(x)\,dx. \quad\text{............(8)}$$

Let $\varphi(\xi) = \int_a^\xi F(x)\,dx$. If when $\xi\to\infty$ the integral $\varphi(\xi)$ becomes and remains greater than any positive number N (or less than any negative number $-N$) the integral (6) is divergent. It may happen, however, that when $\xi\to\infty$ the integral neither converges nor diverges; $\varphi(\xi)$ may tend to no limit and yet be bounded.

For example, if $F(x) = \sin x$, $\varphi(\xi) = \cos a - \cos\xi$ and $|\varphi(\xi)| \leqq 2$ for every ξ.

In this case the integral is said to *oscillate* (finitely). If $\varphi(\xi)$ neither converges nor diverges to $+\infty$ or to $-\infty$ and is not bounded when $\xi\to\infty$, the integral is sometimes said to "oscillate infinitely."

It may happen that the limits in equations (6) and (7) do not exist when ξ and ξ' tend *independently* to infinity, and yet that the limit

$$\mathop{\mathcal{L}}_{\xi\to\infty}\int_{-\xi}^\xi F(x)\,dx$$

is a definite number; in this case the integral is called, as for a finite interval, the Principal Value of the integral. The notations

$$P\int_a^b F(x)\,dx \quad\text{and}\quad P\int_{-\infty}^\infty F(x)\,dx$$

are sometimes used to denote the Principal Values.

If the integral of $F(x)$ over a given interval is convergent, $F(x)$ is said to be *integrable* over the interval.

Absolute Convergence. If the integrals of $F(x)$ and $|F(x)|$ over a given interval are both convergent, the integral of $F(x)$

is said to *converge absolutely* over the interval; or $F(x)$ is said to be *absolutely integrable* over the interval.

For an important contrast between Proper and Improper Integrals in respect of absolute convergence, see § 145, Theorem B.

143. General Conditions for Convergence. In stating the conditions for convergence it is clearly sufficient to consider the convergence at one singular point of an interval (a, b) and at ∞; if c is the singular point, $a<c<b$, the condition for integrability at c must hold whether x tend to c from below or from above, and it will therefore secure brevity without loss of generality to take the singular point to be at an end of the interval. It has to be remembered that we always assume that the proper integral of which the improper is the limit has a definite value—that is, that it exists.

Let a be the only singular point of $F(x)$ in the interval (a, b), and let

$$\varphi(\xi)=\int_\xi^b F(x)\,dx,\quad a<\xi<b.$$

The condition that $\varphi(\xi)$ should tend to a limit when $\xi\to a$ is that, given the arbitrarily small positive number ε, there shall be a positive number δ such that

$$|\varphi(a_1)-\varphi(a_2)|<\varepsilon \text{ if } a<a_1<a_2\leqq a+\delta.$$

Now

$$\varphi(a_1)-\varphi(a_2)=\int_{a_1}^{a_2}F(x)\,dx,$$

and therefore the condition that the integral of $F(x)$ should converge at a, or that $F(x)$ should be integrable at a, is that

$$\left|\int_{a_1}^{a_2}F(x)\,dx\right|<\varepsilon \text{ if } a<a_1<a_2\leqq a+\delta \quad\text{..............(1)}$$

or, what is equivalent, that

$$\mathop{\mathcal{L}}_{a_2\to a}\int_{a_1}^{a_2}F(x)\,dx=0 \text{ if } a<a_1<a_2.$$

If b were the only singular point in an interval (a, b) the condition would be

$$\left|\int_{b_1}^{b_2}F(x)\,dx\right|<\varepsilon \text{ if } b-\delta\leqq b_1<b_2<b,\quad\text{.............(2)}$$

or

$$\mathop{\mathcal{L}}_{b_1\to b}\int_{b_1}^{b_2}F(x)\,dx=0 \text{ if } b_1<b_2<b.$$

In the same way for convergence at ∞ there must be, given ε as usual, a positive number N such that

$$\left|\int_b^c F(x)\,dx\right| < \varepsilon \text{ if } c > b \geqq N, \quad \ldots\ldots\ldots\ldots\ldots\ldots(3)$$

or

$$\mathop{\mathcal{L}}_{b\to\infty}\int_b^c F(x)\,dx = 0 \text{ if } c > b.$$

The integrand always of the same sign. Suppose, for example, that $F(x)$ is not negative for any value of x in the range. In this case the proper integral must be positive (or, at least, cannot be negative) and therefore must tend either to a (finite) limit or to $+\infty$, so that the corresponding improper integral must either converge or diverge and cannot oscillate.

A change of variable, the change being made in the proper integral of which the improper is the limit, will sometimes transform an improper into a proper integral. For example, if $f(x)$ is continuous for $0 \leqq x \leqq \pi/2$, and if $0 \leqq a < 1$ the change from x to y where $x = \sin y$ gives

$$\int_0^a \frac{f(x)\,dx}{\sqrt{(1-x^2)}} = \int_0^{\sin^{-1}a} f(\sin y)\,dy,$$

and therefore

$$\int_0^1 \frac{f(x)\,dx}{\sqrt{(1-x^2)}} = \mathop{\mathcal{L}}_{a\to 1}\int_0^{\sin^{-1}a} f(\sin y)\,dy = \int_0^{\frac{\pi}{2}} f(\sin y)\,dy.$$

Again, an integral over a finite range may be changed into one over an infinite range and *vice versa*. Thus, if $x = e^{-y}$ and $0 < \delta < 1$

$$\int_\delta^1 \log\left(\frac{1}{x}\right)dx = \int_0^{\log\frac{1}{\delta}} ye^{-y}\,dy,$$

and if one of the integrals converges so does the other.

Change of values of the integrand. As in the case of the proper integral, § 110, Theorem II, it is obvious that for the values of $F(x)$, when these values are finite, there may be substituted at any finite number of points in the range any other finite values without changing the value of the integral.

Singular Integral. For brevity, the integrals in (1), (2), (3) will sometimes be called the *singular integrals* at the singular points a, b, ∞; if $a < c < b$ there will be two singular integrals for the point c, and these may be called the *left* and the *right* singular integrals.

144. Special Test. The following test for convergence covers a large class of integrals.

I. *Convergence at a Singular Point.* Let a be the only singular point of an interval (a, b), and suppose there is a neighbourhood $(a, a+\delta)$ of a for which $(x-a)^n F(x) = \varphi(x)$.

(i) If $0<n<1$ and if $\varphi(x)$ is bounded, say $|\varphi(x)|<K$, $F(x)$ is absolutely integrable at a;

(ii) if $n \geqq 1$ and if $\varphi(x)$ is always of the same sign (never zero) in $(a, a+\delta)$, say $\varphi(x)>K_1>0$ or $\varphi(x)<-K_1<0$, where K_1 is a constant, $F(x)$ is not integrable at a.

(i) $0<n<1$, $a<a_1<a_2 \leqq a+\delta$ and $|\varphi(x)| <K$;

$$\left|\int_{a_1}^{a_2} F(x)\,dx\right| \leqq \int_{a_1}^{a_2} |F(x)|\,dx < \int_{a_1}^{a_2} \frac{K\,dx}{(x-a)^n}$$
$$= \frac{K}{1-n}\{(a_2-a)^{1-n} - (a_1-a)^{1-n}\}.$$

Both $(a_2-a)^{1-n}$ and $(a_1-a)^{1-n}$ tend to zero when $a_2 \to a$; therefore both $F(x)$ and $|F(x)|$ are integrable at a.

(ii) Suppose $n \geqq 1$ and that $F(x)$ is positive and $\varphi(x)$ positive, not zero, $\varphi(x)>K_1>0$; then

$$\int_{a_1}^{a_2} F(x)\,dx > \int_{a_1}^{a_2} \frac{K_1\,dx}{(x-a)^n} = \frac{K_1}{n-1}\left\{\frac{1}{(a_1-a)^{n-1}} - \frac{1}{(a_2-a)^{n-1}}\right\},\ n>1$$
$$= K_1 \log\{(a_2-a)/(a_1-a)\},\ n=1.$$

Clearly the integral does not tend to a limit when $a_2 \to a$, and therefore $F(x)$ is not integrable at a. A similar proof holds if $\varphi(x)<-K_1<0$.

If b were the singular point we should put $(b-x)^n F(x) = \varphi(x)$ for the range $b-\delta \leqq b_1 < b_2 < b$.

Ex. 1. Of the integrals

$$\text{(i)}\ \int_a^b \frac{\sin(x-a)\,dx}{(x-a)^{3/2}},\quad \text{(ii)}\ \int_a \frac{1\left(x-a\right)\,dx}{(x-a)^{1/2}}\quad \text{(iii)}\ \int_a^b \frac{\cos(x-a)\,dx}{(x-a)^{3/2}},$$

the first and the second are convergent while the third is not. For (i), note that $(x-a)^{\frac{1}{2}}F(x) \to 1$ when $x \to a$; if n were taken to be 3/2 the integral might seem to be divergent but then $\varphi(a)$ would be zero and the conditions for I (ii) would be violated.

II. *Convergence at* ∞. Suppose that $x^n F(x) = \varphi(x)$ when $x \geqq b$, an arbitrarily large positive number.

(i) If $n>1$ and $|\varphi(x)|<K$, a constant, $F(x)$ converges absolutely at ∞;

(ii) if $n \leqq 1$ and if $\varphi(x)$ is always of the same sign (not zero), say $\varphi(x)>K_1>0$ or $\varphi(x)< -K_1<0$, where K_1 is a constant, the integral of $F(x)$ does not converge at ∞.

(i) If $n>1$, $c>b$ and $|\varphi(x)|<K$,

$$\left|\int_b^c F(x)\,dx\right| \leqq \int_b^c |F(x)|\,dx < \int_b^c \frac{K\,dx}{x^n} = \frac{K}{n-1}\left(\frac{1}{b^{n-1}} - \frac{1}{c^{n-1}}\right).$$

Hence, since $n>1$, $(b^{1-n} - c^{1-n}) \to 0$ when $b \to \infty$, and therefore both $F(x)$ and $|F(x)|$ are integrable at ∞.

(ii) Suppose $n \leqq 1$ and $\varphi(x)$ positive, $\varphi(x)>K_1>0$; then $F(x)$ is positive and

$$\int_b^c F(x)\,dx > \int_b^c \frac{K_1 dx}{x^n} = \frac{K_1}{1-n}(c^{1-n} - b^{1-n}), \quad n<1$$

$$= K_1 \log(c/b), \quad n=1.$$

The integral therefore cannot converge at ∞. A similar proof holds if $\varphi(x)< -K_1<0$.

Ex. 2. The integrals $\int_1^\infty \frac{\sin x\,dx}{x^{3/2}}$ and $\int_0^\infty \frac{\sin x\,dx}{x^{1/2}}$ are both convergent.

For the first integral $n=3/2$ and $\varphi(x)=\sin x$. For the second integral there is convergence at 0 by I (i) above, but the convergence at ∞ cannot be tested by the above rule. However,

$$\int_b^c \frac{\sin x}{x^{1/2}}\,dx = \frac{1}{b^{1/2}}\int_b^\xi \sin x\,dx\,; \quad \left|\int_b^c \frac{\sin x}{x^{1/2}}\,dx\right| \leqq \frac{2}{b^{1/2}},$$

but in this case the convergence is not *absolute*, as may be proved by the method given in the *E.T.* p. 445, Ex. 1.

145. General Theorems. The improper integral has been defined as the limit of a proper integral, and it is therefore necessary to inquire whether certain General Theorems, proved for the proper integral in §§ 109, 111, 112, are valid for the improper integral. The definitions 1 and 2 of § 111 may simply be assumed for the improper integral, and Theorems I, II and VII of § 109 are valid for the improper integral as being either definitions or simple consequences of the definitions. Theorem III, however (and as dependent upon it Theorems IV and V), and also Theorem VI, require modification.

THEOREM A. *If $\varphi(x)$ and $\psi(x)$ are absolutely integrable over an interval so is their product* unless *the functions have the same singular point, in which case the product may or may not be integrable.*

(i) Interval (a, b). Let b be a singular point for $\psi(x)$ but not for $\varphi(x)$ and suppose $|\varphi(x)|<K$ if $b-\delta \leqq x \leqq b$; then, ε and δ having the usual meaning, if $b-\delta \leqq b_1 < b_2 < b$,

$$\int_{b_1}^{b_2} |\varphi(x)\psi(x)|\,dx < K\int_{b_1}^{b_2} |\psi(x)|\,dx < \varepsilon$$

because $|\psi(x)|$ is integrable at b and therefore δ can be chosen so that the inequality is satisfied. From Theorem B below it follows that $\varphi(x)\psi(x)$ is absolutely integrable over any finite interval.

It may again be noted that the integrals used in the proof are proper integrals so that the various inequality theorems may be used.

Cor. If $\varphi(x)$ is bounded and integrable and $\psi(x)$ absolutely integrable so is $\varphi(x)\psi(x)$, because in this case $|\varphi(x)|$ is bounded and integrable.

(ii) Interval (a, ∞). The convergence at ∞ alone needs investigation since the convergence over any finite interval is settled by case (i).

Obviously neither the integral of $|\varphi(x)|$ nor that of $|\psi(x)|$ can converge at ∞ unless $|\varphi(x)|$ and $|\psi(x)|$ are bounded when $x \geqq b$, an arbitrarily large positive number. Suppose $|\varphi(x)|<K'$ when $x \geqq b$; then if $c>b$

$$\int_{b}^{c} |\varphi(x)\psi(x)|\,dx < K'\int_{b}^{c} |\psi(x)|\,dx \to 0 \text{ when } b \to \infty,$$

because $|\psi(x)|$ is integrable at ∞. Hence $|\varphi(x)\psi(x)|$ is integrable, and, from Theorem B, $\varphi(x)\psi(x)$ is absolutely integrable over (a, ∞).

Ex. 1. Let the interval be $(0, 1)$. If $\varphi(x)=(1-x)^{-\frac{1}{2}}$, $\psi(x)=(1-x)^{-2/3}$, both $\varphi(x)$ and $\psi(x)$ are positive and integrable but their product is not; if $\psi(x)=(1-x)^{-1/4}$ the product of $\varphi(x)$ and $\psi(x)$ is integrable.

An extension of this theorem is given below, Theorems E and F.

Theorem B. *If the integral of $|F(x)|$ converges at a singular point or at ∞ so does that of $F(x)$, but the integral of $F(x)$ may converge while that of $|F(x)|$ does not converge.*

Let the integral of $|F(x)|$ converge at b; with the notation of Theorem A we have

$$\left|\int_{b_1}^{b_2} F(x)\,dx\right| \leqq \int_{b_1}^{b_2} |F(x)|\,dx \to 0 \text{ if } b_1 \to b,$$

since $|F(x)|$ is integrable at b. Similarly, if $c>b$,

$$\left|\int_b^c F(x)\,dx\right| \leqq \int_b^c |F(x)|\,dx \to 0 \text{ if } b \to \infty.$$

The first part of the theorem is therefore proved.

It follows that if both $F(x)$ and $|F(x)|$ are integrable

$$\left|\int_a^b F(x)\,dx\right| \leqq \int_a^b |F(x)|\,dx, \quad \left|\int_a^\infty F(x)\,dx\right| \leqq \int_a^\infty |F(x)|\,dx.$$

An example of the truth of the second part has been given in the *Elementary Treatise*, p. 445, when the range of integration is infinite; the following illustrates the case of a finite range.

Ex. 2. Let $F(x)$ be defined for the range (0, 1) as follows, r being any positive integer:

$$F(x) = (-1)^{r-1} r \text{ if } (r+1)^{-1} < x < r^{-1}.$$

The point 0 is a singular point. Now

$$\int_0^1 F(x)dx = \mathop{\mathcal{L}}_{n\to\infty} \left\{ \sum_{r=1}^{n} \frac{(-1)^{r-1}}{r+1} \right\} = 1 - \log 2;$$

but
$$\int_{\frac{1}{n+1}}^1 |F(x)|\,dx = \sum_{r=1}^{n} \frac{1}{r+1} \to \infty \text{ when } n \to \infty.$$

In the case of a proper integral, $|F(x)|$ is always integrable when $F(x)$ is, but for an improper integral this statement is not correct, $F(x)$ may be integrable when $|F(x)|$ is not.

THEOREM C. *An improper integral is a continuous function of its limits.*

If x and $x+h$ are both within the range of integration, we have in the notation of § 112

$$\varphi(x) = \int_a^x F(t)dt, \quad \varphi(x+h) - \varphi(x) = \int_x^{x+h} F(t)dt.$$

When x is not a singular point, δ can be chosen so that the interval $(x-\delta, x+\delta)$ contains no singular point, and therefore when $|h|<\delta$ the function $F(t)$ is bounded and $\varphi(x+h) \to \varphi(x)$ when $h \to 0$, because the proof of § 112 is applicable. On the other hand, if x is a singular point $\varphi(x)$ has a definite value because, by hypothesis, the improper integral exists; next,

$\varphi(x+h) \to \varphi(x)$ when $h \to 0$ by the definition of the convergence of the integral at x.

We can now prove a theorem that is of constant application.

THEOREM D. *If $F(x)$ is integrable over (a, b) and if there is a function $f(x)$ such that* (i) *$f(x)$ is continuous for $a \leqq x \leqq b$ and* (ii) *$F(x)=f'(x)$ except at a singular point, then*

$$\int_a^b F(x)dx = f(b) - f(a).$$

Let c, where $a<c<b$, be the only singular point; then, δ, δ' being as before,

$$\begin{aligned}\int_a^b F(x)\,dx &= \underset{\delta\to 0}{\mathcal{L}}\int_a^{c-\delta} f'(x)\,dx + \underset{\delta'\to 0}{\mathcal{L}}\int_{c+\delta'}^b f'(x)\,dx\\ &= \underset{\delta\to 0}{\mathcal{L}}\{f(c-\delta)-f(a)\} + \underset{\delta'\to 0}{\mathcal{L}}\{f(b)-f(c+\delta')\}\\ &= f(b)-f(a), \text{ since } f(x) \text{ is continuous at } c.\end{aligned}$$

If the range is (a, ∞) we have, if $c>b$,

$$\int_a^\infty F(x)\,dx = \int_a^b f'(x)\,dx + \underset{c\to\infty}{\mathcal{L}}\int_b^c f'(x)\,dx,$$

where we now suppose the conditions to hold for an arbitrarily large interval (a, c). Thus

$$\int_a^\infty F(x)\,dx = -f(a) + \underset{c\to\infty}{\mathcal{L}} f(c) = -f(a) + K$$

if $f(c) \to K$ when $c \to \infty$.

The conditions of Theorem A for the integrability of a product are supplemented by the following Theorems, often called **Abel's Theorem and Dirichlet's Theorem** respectively.

THEOREM E, or, Abel's Theorem. *If $\varphi(x)$ is bounded and monotonic and if $\psi(x)$ is integrable, whether the range of integration is* (i) *finite (a, b) or* (ii) *infinite (a, ∞), their product is integrable.*

(i) $\varphi(x)$ is integrable since it is bounded and monotonic. If b is a singular point for $\psi(x)$ and if $b-\delta \leqq b_1 < b_2 < b$, we have by the Second Theorem of Mean Value for proper integrals $(b_1 \leqq \xi \leqq b_2)$

$$\int_{b_1}^{b_2}\varphi(x)\,\psi(x)\,dx = \varphi(b_1+0)\int_{b_1}^{\xi}\psi(x)\,dx + \varphi(b_2-0)\int_{\xi}^{b_2}\psi(x)\,dx$$

Each of the integrals on the right tends to zero when $b_1 \to b$, since $\psi(x)$ is integrable at b while $\varphi(b_1+0)$ and $\varphi(b_2-0)$ are finite; the integral of the product is therefore convergent at b, so that the product is integrable over any finite range.

(ii) Take b so large that all the singular points of $\psi(x)$ lie within (a, b); then, if $c>b$ and $b \leqq \xi \leqq c$,

$$\int_b^c \varphi(x)\,\psi(x)\,dx = \varphi(b+0)\int_b^\xi \psi(x)\,dx + \varphi(c-0)\int_\xi^c \psi(x)\,dx\,;$$

as before, it follows that the product is integrable over (b, ∞) and therefore by (i) over (a, ∞).

Cor. The product is absolutely integrable if $\psi(x)$ is so.

THEOREM F, or, Dirichlet's Theorem. *If $\varphi(x)$ is monotonic and tends to zero when $x\to\infty$, and if the integral of $\psi(x)$ converges over an arbitrarily large interval (a, b) but oscillates (finitely) at ∞, the integral*

$$\int_a^\infty \varphi(x)\psi(x)\,dx$$

is convergent.

By Theorem E the product $\varphi(x)\,\psi(x)$ is integrable over (a, b). If $c>b$ and $b \leqq \xi \leqq c$ we have

$$\left|\int_b^c \varphi(x)\,\psi(x)\,dx\right| = \left|\varphi(b+0)\int_b^\xi \psi(x)\,dx\right| < K\,|\varphi(b+0)|$$

since the integral of $\psi(x)$ oscillates finitely. But $\varphi(b+0)\to 0$ when $b\to\infty$ and K is finite so that the integral of the product converges at ∞.

Theorems E and F give useful tests for the convergence of an integral. Abel's Theorem shows that a convergent integral remains convergent when the integrand is multiplied by a bounded monotonic factor, while Dirichlet's shows that an oscillating integral may be made convergent by multiplying the integrand by a monotonic factor which tends to zero when x tends to infinity.

Ex. 3. Discuss Dirichlet's Theorem for a finite interval (a, b).

The student should now have little difficulty in extending the Fundamental Inequality Theorem and the two Theorems of Mean Value to improper integrals; a sketch of the proofs will therefore be sufficient.

If a is the only singular point in (a, b) and if $F(x)$ is integrable over (a, b) and not negative, then, $a<a_2$,

$$\int_{a_2}^{b} F(x)dx \geqq 0, \quad \int_{a}^{b} F(x)dx = \mathop{\mathcal{L}}_{a_2 \to a} \int_{a_2}^{b} F(x)dx \geqq 0.$$

Similarly, if $F(x)$ is integrable over (a, ∞), the integral is not negative.

The First Theorem of Mean Value follows at once. In the notation of § 111 put $\varphi(x)\{\psi(x)-g\}$ and then $\varphi(x)\{G-\psi(x)\}$ for $F(x)$; these products are integrable by Theorem A, and the method (*E.T.* § 124) applies.

For the Second Theorem of Mean Value, with the notation of § 111, $\psi(x)$ being integrable over (a, b) so is $\varphi(x)\psi(x)$ by Abel's Theorem. Now suppose that c, where $a<c<b$, is the only singularity of $\psi(x)$ in (a, b) and enclose c in the interval $(c-\delta, c+\delta)$ where δ is positive and arbitrarily small; since the integrals of $\psi(x)$ and $\varphi(x)\varphi(x)$ exist, the interval $(c-\delta, c+\delta)$ instead of $(c-\delta, c+\delta')$ may be taken. Choose $\psi_1(x)$ so that

$$\psi_1(x) = \psi(x) \text{ if } a \leqq x \leqq c-\delta \text{ or } c+\delta \leqq x \leqq b,$$

but $\psi_1(x) = 0$ if $c-\delta<x<c+\delta$.

Let
$$I = \int_a^b \varphi(x)\psi(x)dx, \quad I_1 = \int_a^b \varphi(x)\psi_1(x)dx,$$

and
$$f(x) = \int_a^x \psi(t)dt, \qquad f_1(x) = \int_a^x \psi_1(t)dt.$$

All these integrals exist and $I_1 \to I$, $f_1(x) \to f(x)$ when $\delta \to 0$; for, by the definition of $\psi_1(x)$,

$$|I - I_1| = \left| \int_{c-\delta}^{c+\delta} \varphi(x)\psi(x)dx \right| = \eta_\delta, \quad |f(x) - f_1(x)| \leqq \left| \int_{c-\delta}^{c+\delta} \psi(t)dt \right| = \eta'_\delta,$$

and, since the integrals converge at c, both η_δ and η'_δ tend to zero when δ tends to zero.

The Mean Value Theorem holds for the integral I_1 since $\psi_1(x)$ is finite; therefore if g_1 and G_1 are the minimum and maximum value of $f_1(x)$ for $a \leqq x \leqq b$ we have

$$I_1 - g_1\varphi(a+0) \geqq 0, \quad G_1\varphi(a+0) - I_1 \geqq 0.$$

Hence, since $g_1 \to g$ and $G_1 \to G$ when $\delta \to 0$, we find

$$g\varphi(a+0) \leqq \int_a^b \varphi(x)\psi(x)dx \leqq G\varphi(a+0).$$

The theorem, being now proved when there is one singularity,

can be extended to the case of a second singularity c' by the method just used and then, in succession, to the case of m singularities. A similar method is obviously applicable when the range of integration is infinite (a, ∞). The usual form

$$\int_a^b \varphi(x)\,\psi(x)\,dx = \varphi(a+0)\int_a^\xi \psi(x)\,dx$$

is deduced as before, b being finite or $+\infty$.

Transformations of the Improper Integral. In practice it is advisable to carry out a change in the variable of integration by operating on the proper integral and passing to the limit; if the precautions required for operating on the proper integral are observed there will be, as a rule, little difficulty in completing the transformation, so that little or nothing is to be gained by elaborating any special rules for the improper integral.

A similar observation is applicable to the rule of integration by parts. The student's "common sense" may be left with some range of operations on which to exercise itself.

Ex. 4. $\displaystyle\int_0^1 x^{m-1}(1-x)^{n-1}dx = \int_0^\infty \frac{y^{n-1}dy}{(1+y)^{m+n}},\ m>0,\ n>0.$

If $0<m<1$ and $0<n<1$, the integral

$$B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx$$

is an improper integral. Let λ and μ be small positive numbers; the integral

$$\int_\lambda^{1-\mu} x^{m-1}(1-x)^{n-1}dx$$

is a proper integral, and if $x = 1/(1+y)$ the integral becomes

$$\int_{\frac{\mu}{1-\mu}}^{\frac{1}{\lambda}-1} \frac{y^{n-1}dy}{(1+y)^{m+n}}.$$

The limit of this integral when $\lambda\to 0$ and $\mu\to 0$ is the convergent integral

$$\int_0^\infty \frac{y^{n-1}dy}{(1+y)^{m+n}}$$

so that $\displaystyle B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx = \int_0^\infty \frac{y^{n-1}dy}{(1+y)^{m+n}}.$

146. Worked Examples. Additional Tests. The Special Test, given in § 144, and the Tests of Abel and Dirichlet, given in § 145, are sufficient for determining the convergence of large classes of integrals, but, as in the analogous theory of

the convergence of series, there is no set of Tests that cover all cases. In this article examples will be worked to illustrate the use of the Tests mentioned, and some additional methods of considering the convergence of improper integrals will be indicated. The student is reminded of the assumptions stated in § 141, *Note*.

Such limits as $\underset{x \to 0}{\mathcal{L}} x^n \log x$, $n > 0$ and $\underset{x \to \infty}{\mathcal{L}} x^n e^{-x}$ (*E.T.* p. 99) are often required, and elementary transformations of a very simple kind are frequently effective. For example, if $0 < \alpha < \beta$, let $\sin x$ be put in the form $(\sin x/x)x$, taking the value of $(\sin x/x)$ for $x = 0$ to be 1; then

$$\int_\alpha^\beta \log \sin x \, dx = \int_\alpha^\beta \log\left(\frac{\sin x}{x}\right) dx + \int_\alpha^\beta \log x \, dx \quad \ldots\ldots\ldots (a)$$

Now $$\int \log x \, dx = x \log x - x,$$

and therefore the integral of $\log \sin x$ converges at 0, because each of the two integrals on the right of equation (a) tends to 0 when $\beta \to 0$.

Again, if $0 < x < 1$, $x^{-1} \log(1+x) = 1 - \frac{1}{2}x + \ldots$ so that the integral of $x^{-1} \log(1+x)$ converges at 0; the singularity in this case is "removable," as in that of $(\sin x/x)$, by defining the function for the value $x = 0$ to be the limit for $x \to 0$. (See *E.T.* p. 418; also § 29 above.)

See also the remarks about change of variable in § 143.

Ex. 1. The integral $B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1} dx$ is convergent (absolutely) if, and only if, $m > 0$ and $n > 0$.

So far as the question of convergence is concerned we may take the integrand to be x^{m-1} near the lower limit and $(1-x)^{n-1}$ near the upper limit; because near these limits the integrand is of the forms $A x^{m-1}$ and $B(1-x)^{n-1}$ respectively where A and B differ but little from unity and the convergence is not affected by the particular values of A and B so long as these are finite. The Test of § 144 then gives the relations $m > 0$ and $n > 0$.

Ex. 2. The integral $\int_a \frac{f_m(x)}{f_n(x)} dx$, where $f_m(x)$ and $f_n(x)$ are polynomials of degrees m and n respectively, is (absolutely) convergent if $n \geqq m+2$, (i) provided a is greater than the greatest root of $f_n(x) = 0$, (ii) for every value of a, including $-\infty$, if the equation $f_n(x) = 0$ has no real roots; it diverges if $n < m+2$ for every value of a.

The integrand is of the form $Ax^{m-n}\{1+\eta_x\}$ where $\eta_x \to 0$ if $x \to \infty$; the result then follows by § 144.

Ex. 3. Of the following integrals,

$$\int_a^b \frac{Ax^2+Bx+C}{\sqrt{\{(x-a)(b-x)\}}}\,dx, \quad \int_a^b \frac{A\,dx}{(x-a)\sqrt{(b-x)}}, \quad \int_0^4 \frac{dx}{\sqrt[3]{\{(x-1)(x-2)^2\}}}$$

the first and third converge while the second diverges. Show that the first is reducible to a proper integral by the substitution

$$x = a\cos^2\theta + b\sin^2\theta.$$

Ex. 4. The integral $\int_0^{\frac{\pi}{2}} (\sin x)^{m-1}(\cos x)^{n-1}\,dx$ converges if, and only if, $m>0$, $n>0$.

Here $(\sin x)^{m-1} = (\sin x/x)^{m-1} \,.\, x^{m-1}$ so that, near the lower limit the integrand may be taken as x^{m-1}; similarly, near the upper limit the integrand may be taken to be $\left(\frac{\pi}{2}-x\right)^{n-1}$. Apply § 144.

If $0<\alpha<\beta<\frac{\pi}{2}$ and $\sin x = y^{\frac{1}{2}}$ we have

$$\int_\alpha^\beta (\sin x)^{m-1}(\cos x)^{n-1}\,dx = \tfrac{1}{2}\int_{\sin^2\alpha}^{\sin^2\beta} y^{\frac{m}{2}-1}(1-y)^{\frac{n}{2}-1}\,dy\,;$$

let $\alpha \to 0$ and $\beta \to \frac{\pi}{2}$; the given integral is $\frac{1}{2}B\left(\frac{m}{2}, \frac{n}{2}\right)$ by Ex. 1.

Ex. 5. The integral $\int_0^{\frac{\pi}{2}} \frac{x^m\,dx}{(\sin x)^n}$ converges if, and only if, $n<m+1$ and $\int_0^{\frac{\pi}{2}} \frac{\log(\sin x)}{(\sin x)^n}\,dx$ converges if $n<1$.

Note that the first integrand $= x^{m-n}(x/\sin x)^n$.

Ex. 6. The integral $\int_0^\infty \frac{x^{p-1}\,dx}{1+x}$ converges if, and only if, $0<p<1$.

Apply the method of Examples 1 and 2.

Ex. 7. The integral $\int_0^\infty \frac{\sin(ax+b)}{x^n}\,dx$ converges if $0<n<1$, $a\neq 0$, $b\neq 0$, but if $0<n<2$, $a\neq 0$, $b=0$.

By Dirichlet's Test the integral converges at ∞ if $n>0$; for convergence at 0 apply § 144.

Ex. 8. The integral $\int_0^1 (x^p+x^{-p})\log(1+x)\frac{dx}{x}$ converges if $-1<p<1$.

Ex. 9. Prove that $\int_0^\infty \frac{x^{p-1}-x^{-p}}{1-x}\,dx = 2\int_0^1 \frac{x^{p-1}-x^{-p}}{1-x}\,dx$ if $0<p<1$.

The value of the integrand for $x=1$ may be taken as $(1-2p)$.

Ex. 10. Prove that, if 0 is the only singular point for $F(x)$ in $(0, a)$,

$$\text{(i)} \int_0^a F(x)\,dx = \int_0^a F(a-x)\,dx\,;$$

then, as in *E.T.* p. 332, Example 6, show that

$$\text{(ii)} \int_0^{\frac{\pi}{2}} \log \sin x\,dx = -\frac{\pi}{2}\log 2.$$

For (i), if $0<\delta<a$, $\int_\delta^a F(x)\,dx = \int_0^{a-\delta} F(a-y)\,dy$; then let $\delta \to 0$.

Ex. 11. Show that $\int_0^1 \log \Gamma(x)\,dx$ is convergent and then, by using the relation (§ 96, (4)) $\Gamma(x)\Gamma(1-x)=\pi/\sin \pi x$ and the results in Ex. 10, prove that

$$\int_0^1 \log \Gamma(x)\,dx = \tfrac{1}{2}\log(2\pi).$$

Since $\Gamma(x)=\Gamma(x+1)/x$ and $\log\Gamma(x)=\log\Gamma(x+1)-\log x$ the integral converges at 0; then, as in *E.T.* p. 332, Example 6,

$$\int_0^1 \log \Gamma(x)dx = \int_0^1 \log\Gamma(1-x)\,dx = \tfrac{1}{2}\int_0^1 \log[\Gamma(x)\Gamma(1-x)]\,dx,$$

and therefore $$\int_0^1 \log\Gamma(x)\,dx = \tfrac{1}{2}\int_0^1 \log[\pi/\sin \pi x]\,dx, \text{ etc.}$$

Ex. 12. If $u=\int_x^{x+1} \log\Gamma(t)\,dt$, $x \geqq 0$, prove that $du/dx=\log x$ and deduce that

$$u = x\log x - x + \tfrac{1}{2}\log(2\pi).$$

If $x<a<x+1$, $u=\int_a^{x+1} \log\Gamma(t)\,dt - \int_a^x \log\Gamma(t)\,dt,$

and therefore $$\frac{du}{dx} = \log\Gamma(x+1) - \log\Gamma(x) = \log x.$$

Now integrate and apply Ex. 11 to determine the constant of integration.

Ex. 13. The integral $\int_0^1 \cos x \,\log x\,dx$ converges but the integral $\int_1^\xi \cos x\,\log x\,dx$ oscillates when $\xi\to\infty$.

A simple but useful Test, called from analogy with a test for series (*E.T.* p. 380) the **Comparison Test**, is derived as follows:

Let $$\varphi(\xi)=\int_b^\xi F(x)dx, \quad \psi(\xi)=\int_b^\xi f(x)dx,$$

where b is an arbitrarily large (positive) number. If $F(x)$ is positive for $x \geqq b$ the function $\varphi(\xi)$ is positive, monotonic and increasing, so that if $\varphi(\xi)$ is bounded, say $\varphi(\xi)<K$, a positive constant, for $\xi>b$ it tends to a limit and the integral of $F(x)$ converges absolutely at ∞, while if $\varphi(\xi)$ is not bounded for $\xi>b$ the integral of $F(x)$ diverges at ∞.

Again, if $F(x) < f(x)$ for $x \geqq b$ $(f(x) > 0)$, and if the integral of $f(x)$ converges at ∞, so that $\psi(\xi)$ is bounded for $\xi > b$, then $\varphi(\xi) < \psi(\xi)$, and therefore the integral of $F(x)$ converges (absolutely) at ∞ if that of $f(x)$ does so. Similarly, when $F(x) > f(x)$ for $x \geqq b$, it is seen that if the integral of $f(x)$ diverges at ∞ so does that of $F(x)$.

A similar test for convergence at the ends a and b of an interval (a, b) may be derived by considering the integrals

$$\varphi_1(\xi) = \int_\xi^b F(x)\,dx, \quad \varphi_2(\xi) = \int_a^\xi F(x)\,dx.$$

The Comparison Test applies only to absolute convergence.

Ex. 14. The integral $\int_0^\infty e^{-kx}x^{n-1}dx = \Gamma(n)/k^n, \quad \begin{matrix} k > 0 \\ n > 0. \end{matrix}$

(i) Convergence at ∞. If $k > 0$ and p is any positive integer

$$e^{kx} > (kx)^p/p!\,; \quad e^{-kx}x^{n-1} < K/x^2$$

where p is so chosen that $p - n > 1$ and K is a constant $(k^{-p}p!)$. In the comparison test let $f(x) = K/x^2$ and, by § 144, the integral converges at ∞ for every value of n.

If $k = 0$, the integral converges at ∞ if n is negative (not zero), while if $k < 0$ the integral obviously diverges at ∞.

(ii) Convergence at 0. By § 144 the integral converges at the lower limit if, and only if, $n > 0$.

Thus the integral converges if $k > 0$ and $n > 0$ and diverges in all other cases.

From this example a useful *Comparison Function* is derived. Put $n+1$ for n and let $f(x) = Ae^{-kx}x^n$, $A > 0$; the integral of $f(x)$ converges at ∞, if $k > 0$, for every n but diverges at ∞, if $k < 0$, for every n. If $k = 0$ the test of § 144 may be used.

Ex. 15. If $f(x) = Ax^{-k-1}(\log x)^n$, $A > 0$, $x > b$ (arbitrarily large), the integral of $f(x)$ converges at ∞, if $k > 0$, for every n but diverges at ∞, if $k < 0$, for every n.

Let $\log x = t$; then

$$\int_b^\xi f(x)\,dx = \int_B^T Ae^{-kt}t^n\,dt, \quad \begin{matrix} T = \log \xi \\ B = \log b. \end{matrix}$$

When $\xi \to \infty$ so does T and the result follows from the comparison function $Ae^{-kt}t^n$. Thus, by the change of variable, another comparison function is obtained.

Ex. 16. If $f(x) = Ax^{k-1}[\log(1/x)]^n$, $A > 0$, $0 < x \leqq 1$, the integral of $f(x)$ converges at 0, if $k > 0$, for every n, but diverges at 0, if $k < 0$, for every n; if $k = 0$, the integral converges at 0, if, and only if, $n < -1$. The integral converges at 1 if, and only if, $n > -1$.

Let $x=e^{-t}$; then if $0<\alpha<\beta<1$,

$$\int_{\alpha}^{\beta} f(x)\,dx=\int_{b}^{a} Ae^{-kt}t^{n}\,dt, \quad \begin{array}{l} a=\log(1/\alpha) \\ b=\log(1/\beta),\ a>b. \end{array}$$

When $\beta\to 0$, $b\to\infty$ and the result for $k\gtreqless 0$ follows, as before, for convergence at 0. If $k=0$ the test of § 144 applies for convergence at 0; the same test applies for convergence at 1 for every value of k.

Ex. 17. The integral $\int_0^\infty e^{-x}x^{n-1}(\log x)^m\,dx$ converges (absolutely) if $n>0$ and m any positive integer.

If $0<x<1$, the integrand is less than $x^{n-1}(-\log x)^m$, numerically, while for large values of x the integrand is less than $e^{-x}x^{n-1+m}$.

Now apply the results of Example 16 for convergence at 0.

Ex. 18. If $a>1$, the integral $\int_a^\infty \frac{dx}{x(\log x)^m}$ converges if $m>1$ but diverges if $m\leqq 1$.

Ex. 19. Let lx, l^2x, l^3x, ... denote $\log x$, $\log(\log x)$, $\log[\log(\log x)]$, ... Show that, when x is large enough to make $l^m x$ positive, the integral of $f(x)$ where

$$f(x)=1/[x\,.\,lx\,.\,l^2x\ \dots\ l^{m-1}x\,.\,(l^m x)^k]$$

converges at ∞ if $k>1$, but diverges at ∞ if $k\leqq 1$.

Let $x=e^{x_1}$ so that $lx=x_1$, $l^2x=lx_1$, ... $l^m x=l^{m-1}x_1$; then if $c>b$,

$$\int_b^c f(x)\,dx=\int_{b_1}^{c_1}\frac{dx_1}{x_1\,.\,lx_1\,.\,l^2x_1\ \dots\ l^{m-2}x_1\,.\,(l^{m-1}x_1)^k}, \quad \begin{array}{l} c_1=\log c \\ b_1=\log b. \end{array}$$

Next let $x_1=e^{x_2}$, then $x_2=e^{x_3}$ and so on; the integral is thus reduced to

$$\int_b^c f(x)\,dx=\int_{b_m}^{c_m}\frac{dx_m}{x_m^k}, \quad \begin{array}{l} c_m=l^m c \\ b_n=l^m b. \end{array}$$

Now apply § 144.

Ex. 20. (i) If a^2 and b^2 are both different from zero the integral

$$\int_0^\infty x^n e^{-a^2x^2-b^2/x^2}\,dx$$

converges (absolutely) for every value of n.

(ii) $\int_0^\infty \frac{x^n\log x\,dx}{(1+x^2)^{n+1}}=0$ if $n>-1$.

Deduce that, if $a>0$,

$$\int_0^\infty \frac{x^n\log x\,dx}{(a^2+x^2)^{n+1}}=\frac{1}{2}\frac{\log a}{a^{n+1}}B\left(\frac{n+1}{2},\frac{n+1}{2}\right).$$

When the range of integration is not bounded it is often useful to express the integral as a series; the following examples illustrate the method.

Ex. 21. The integral $\int_0^\infty \frac{\sin x}{\sqrt{x}}\,dx=2\int_0^\infty \sin(x^2)\,dx$ converges conditionally.

If $n\pi \leqq \xi < (n+1)\pi$, divide the interval $(0, \xi)$ into the partial intervals $(0, \pi)$, $(\pi, 2\pi)$ ... $[(n-1)\pi, n\pi]$, $(n\pi, \xi)$. Now

$$\int_{r\pi}^{(r+1)\pi} \frac{\sin x}{\sqrt{x}}\, dx = (-1)^r \int_0^\pi \frac{\sin y\, dy}{\sqrt{(r\pi+y)}}, \quad x = r\pi + y;$$

therefore

$$\int_0^\xi \frac{\sin x\, dx}{\sqrt{x}} = \sum_{r=0}^{n-1} (-1)^r \int_0^\pi \frac{\sin y\, dy}{\sqrt{(r\pi+y)}} + R_n = \sum_{r=0}^{n-1} u_r + R_n,$$

where

$$|R_n| = \left|\int_0^{\xi - n\pi} \frac{(-1)^n \sin y}{\sqrt{(n\pi+y)}}\, dy \right| \quad \frac{}{\sqrt{(n\pi)}}.$$

When $\xi \to \infty$, $R_n \to 0$ and the series Σu_r is a convergent alternating series so that the integral converges, but not absolutely.

The integral is equal to $(\pi/2)^{\frac{1}{2}}$ (*E.T.* p. 471).

Ex. 22. If $\alpha \geqq 0$ and $\beta > 0$, the integral $\int_0^\infty \frac{x^\alpha\, dx}{1 + x^\beta \sin^2 x}$ converges absolutely if $\beta > 2(\alpha+1)$ but diverges if $\beta \leqq 2(\alpha+1)$.

The integrand is never negative, so that the integral cannot oscillate. Now if n is zero or a positive integer and $n\pi \leqq x \leqq (n+1)\pi$,

$$\frac{(n\pi)^\alpha}{1 + (n+1)^\beta \pi^\beta \sin^2 x} \leqq \frac{x^\alpha}{1 + x^\beta \sin^2 x} \leqq \frac{(n+1)^\alpha \pi^\alpha}{1 + (n\pi)^\beta \sin^2 x}.$$

But if A and B are positive

$$\int_0^\pi \frac{A\, dx}{1 + B\sin^2 x} = 2\int_0^{\frac{\pi}{2}} \frac{A \operatorname{cosec}^2 x\, dx}{\cot^2 x + 1 + B} = \frac{\pi A}{\sqrt{(1+B)}},$$

and therefore, from the inequalities,

$$u_n = \frac{n^\alpha \pi^{\alpha+1}}{\sqrt{\{1 + (n+1)^\beta \pi^\beta\}}} < \int_{n\pi}^{(n+1)\pi} \frac{x^\alpha dx}{1 + x^\beta \sin^2 x} < \frac{(n+1)^\alpha \pi^{\alpha+1}}{\sqrt{(1 + n^\beta \pi^\beta)}} = v_n.$$

Each of the series Σu_n and Σv_n converges (absolutely) if $\beta > 2(\alpha+1)$ but diverges if $\beta \leqq 2(\alpha+1)$ because Σu_n and Σv_n behave like $\Sigma 1/n^{\frac{1}{2}\beta - \alpha}$. Hence the integral behaves as stated.

Ex. 23. As particular cases of Example 22, prove that the integrals

$$\int_0^\infty \frac{dx}{1 + x^4 \sin^2 x} \quad \text{and} \quad \int_0^\infty \frac{x\, dx}{1 + x^6 \sin^2 x}$$

are convergent.

Prove by a similar method that the integrals

$$\int_0^\infty \frac{x\, dx}{1 + x^3 |\sin x|} \quad \text{and} \quad \int_0^\infty \frac{x\, dx}{1 + x^2 |\sin x|}$$

are respectively convergent and divergent. These are particular cases of the integral

$$\int_0^\infty \frac{x^\beta dx}{1 + x^\alpha |\sin x|}, \quad \alpha > 0,\ \beta > 0,$$

(See Hardy, *Messenger of Mathematics*, XXXI, Note VIII) which converges if $\alpha > \beta + 1$ and diverges if $\alpha \leqq \beta + 1$. [The series of Notes contains much instructive analysis.]

Ex. 24. Show by dividing the range into partial intervals, as in *E.T.* p. 449, that

$$\text{(i)}\quad \int_0^\infty F_1(\sin x)\frac{dx}{x}=\int_0^{\frac{\pi}{2}} F_1(\sin x)\frac{dx}{\sin x},$$

if $F_1(u)$ is an odd function of u, and

$$\text{(ii)}\quad \int_0^\infty F_2(\sin x)\frac{dx}{x^2}=\int_0^{\frac{\pi}{2}} F_2(\sin x)\frac{dx}{(\sin x)^2},$$

if $F_2(u)$ is an even function of u, it being understood that the integrals converge.

If $0<\delta\leqq x\leqq\frac{\pi}{2}$, an expression for $\operatorname{cosec}^2 x$ in terms of partial fractions is given in Exercises XII, Ex. 3, (iii) by substituting x for πx; namely

$$\frac{1}{\sin^2 x}=\frac{1}{x^2}+\sum_{n=1}\left[\frac{1}{(n\pi+x)^2}+\frac{1}{(n\pi-x)^2}\right].$$

Of the methods applicable to particular types of integrals one of the most interesting is a method of treating Frullani's Integral (*E.T.* p. 480, Ex. 22). An important article by Hardy, *A Generalisation of Frullani's Integral*, *Messenger of Mathematics*, XXXIV, pp. 11-18, p. 102, should be consulted; also Bromwich, *Inf. Ser.* 2nd Ed. p. 479.

Ex. 25. **Frullani's Integral.** $\displaystyle\int_0^\infty \frac{\varphi(ax)-\varphi(bx)}{x}dx,\ \begin{matrix}a>0\\ b>0\end{matrix}.$

Suppose that $\varphi(x)$ is integrable over an arbitrarily large interval (λ, μ) where $\lambda>0$ and that the integral of $\varphi(x)$ either converges or oscillates finitely at ∞; by Abel's or Dirichlet's Test $\varphi(x)/x$ and therefore also, since a and b are positive, $\varphi(ax)/x$ and $\varphi(bx)/x$ are integrable over (λ, ∞).

Now, putting t in turn for ax and bx, we find

$$\int_\lambda^\infty \frac{\varphi(ax)}{x}dx=\int_{\lambda a}^\infty \frac{\varphi(t)}{t}dt,\quad \int_\lambda^\infty \frac{\varphi(bx)}{x}dx=\int_{\lambda b}^\infty \frac{\varphi(t)}{t}dt,$$

and therefore

$$\int_\lambda^\infty \frac{\varphi(ax)-\varphi(bx)}{x}dx=\int_{\lambda a}^{\lambda b}\frac{\varphi(t)}{t}dt=\int_a^b \frac{\varphi(\lambda x)}{x}dx.$$

(i) Suppose that $\varphi(x)\to N$, a definite number, when $x\to 0$. The numbers a and b are positive and finite so that λ can be chosen so small that $|\varphi(\lambda x)-N|$ will be arbitrarily small for the range $a\leqq x\leqq b$; the limit for $\lambda\to 0$ may therefore be found by putting N for $\varphi(\lambda x)$ in the integral. Hence

$$\int_0^\infty \frac{\varphi(ax)-\varphi(bx)}{x}dx=\mathop{\mathcal{L}}_{\lambda\to 0}\int_a^b \frac{\varphi(\lambda x)}{x}dx=\int_a^b \frac{N}{x}dx=N\log\frac{b}{a}.$$

(ii) Suppose that $\varphi(x) \to M$, a definite number, when $x \to \infty$; in this case the integral of $\varphi(x)$ diverges at ∞ unless $M=0$ and the preceding work requires modification. But, if $\varphi(x)$ is integrable over the arbitrary interval (λ, μ), we have

$$\int_{\lambda}^{\mu} \frac{\varphi(ax)-\varphi(bx)}{x}\,dx = \int_{\lambda a}^{\mu a} \frac{\varphi(t)}{t}\,dt - \int_{\lambda b}^{\mu b} \frac{\varphi(t)}{t}\,dt = \int_{\lambda a}^{\lambda b} \frac{\varphi(t)}{t}\,dt - \int_{\mu a}^{\mu b} \frac{\varphi(t)}{t}\,dt$$

so that

$$\int_{\lambda}^{\mu} \frac{\varphi(ax)-\varphi(bx)}{x}\,dx = \int_{a}^{b} \frac{\varphi(\lambda x)-\varphi(\mu x)}{x}\,dx.$$

Hence, if $\varphi(x) \to N$ when $x \to 0$ and $\varphi(x) \to M$ when $x \to \infty$, we find by letting $\lambda \to 0$ and $\mu \to \infty$ that

$$\int_{0}^{\infty} \frac{\varphi(ax)-\varphi(bx)}{x}\,dx = (N-M)\int_{a}^{b} \frac{dx}{x} = (N-M)\log\frac{b}{a}.$$

Ex. 26. If $a>0$, $b>0$, $n>0$, deduce from Ex. 25:

(i) $\displaystyle\int_{0}^{\infty} \frac{e^{-ax}-e^{-bx}}{x}\,dx = \log\frac{b}{a}$; (ii) $\displaystyle\int_{0} \frac{e \quad - e}{x}\,dx = \log n$;

(iii) $\displaystyle\int_{0}^{\infty} \frac{\cos ax - \cos bx}{x}\,dx = \log\frac{b}{a}$;

(iv) $\displaystyle\int_{0}^{\infty} \frac{e^{-ax}-e^{-bx}-(b-a)xe^{-bx}}{x^2}\,dx = b-a-a\log\frac{b}{a}$.

Ex. 27. If $a, b, c, \ldots k$ are positive, and if the constants $A, B, C, \ldots K$ satisfy the equations

(i) $\Sigma A = A+B+\ldots+K=0$; (ii) $\Sigma Aa = Aa+Bb+\ldots+Kk=0$,

prove that

$$\int_{0}^{\infty} \{\Sigma Ae^{-ax}\}\frac{dx}{x^2} = \Sigma Aa\log a$$

where $\qquad \Sigma Ae^{-ax} = Ae^{-ax}+Be^{-bx}+\ldots+Ke^{-kx}$

and $\qquad \Sigma Aa\log a = Aa\log a + Bb\log b+\ldots+Kk\log k.$

Conditions (i) and (ii) show that the integral converges at $x=0$. The result may be proved by integration by parts. Denote the integrand by $F(x)$; then, $0<\lambda<\mu$,

$$\int_{\lambda}^{\mu} F(x)\,dx = \left[-\frac{\Sigma Ae^{-ax}}{x}\right]_{\lambda}^{\mu} + \int_{\lambda}^{\mu} \frac{-\Sigma Aa\,e^{-ax}}{x}\,dx.$$

By equations (i) and (ii), the integrated part tends to 0 when $\lambda \to 0$; also it tends to 0 when $\mu \to \infty$ since $a, b, \ldots k$ are positive. Again, by equation (ii),

$$-\Sigma Aa\,e^{-ax} = \Sigma Aa\,e^{-x} - \Sigma Aa\,e^{-ax} = \Sigma Aa(e^{-x}-e^{-ax}),$$

and now, if $\lambda \to 0$ and $\mu \to \infty$, the result follows by Ex. 26 (ii).

Ex. 28. Suppose that $\varphi(x)$ and $\varphi'(x)$ are continuous and integrable over an arbitrarily large interval (λ, μ), $\lambda>0$, and that the integrals of $\varphi(x)$ and $\varphi'(x)$ either converge or oscillate finitely at ∞, the functions $\varphi(x)$ and $\varphi'(x)$ being continuous and expressible by Maclaurin's Theorem near $x=0$.

If $\Sigma A = 0$ and $\Sigma Aa = 0$, prove that

$$\int_0^\infty \{\Sigma A\varphi(ax)\}\frac{dx}{x^2} = -\varphi'(0)\Sigma Aa \log a.$$

Proceed as in Ex. 27. Note that, since $\Sigma Aa = 0$,

$$\Sigma Aa\,\varphi'(ax) = -\Sigma Aa[\varphi'(x) - \varphi'(ax)],$$

and
$$\int_0^\infty \frac{\varphi'(x) - \varphi'(ax)}{x}\,dx = \varphi'(0)\log a.$$

For example, let $\varphi(x) = \cos x$.

For further developments consult the paper by Hardy in the *Messenger*.

147. Complex Functions of a Real Variable. A detailed discussion of complex functions of a real variable is outside the scope of this book, but a few of the more important cases may be noted.

Notation. When $c = \alpha + i\beta$, where α and β are real, the real part α will often be denoted by $R(c)$ and the coefficient β of the imaginary part $i\beta$ by $I(c)$ or $R(c/i)$.

Ex. 1. $\int_a^\infty e^{-cx}\,dx = \frac{e^{-ca}}{c}$, a real, $R(c) > 0$.

Let $c = \alpha + i\beta$; then $|e^{-cx}| = e^{-\alpha x}$ and

$$\left|\int_b^\infty e^{-cx}\,dx\right| \leqq \int_b^\infty e^{-\alpha x}\,dx = \frac{e^{-\alpha b}}{\alpha}$$

so that this integral tends to zero when $b \to \infty$, since $\alpha = R(c) > 0$.

But
$$\int_a^b e^{-cx}\,dx = \frac{e^{-ca}}{c} - \frac{e^{-cb}}{c}$$

and therefore
$$\int_a^\infty e^{-cx}\,dx = \mathop{\mathcal{L}}_{b\to\infty} \int_a^b e^{-cx}\,dx = \frac{e^{-ca}}{c}.$$

The integral is therefore evaluated by the usual rule when $R(c) > 0$.

Ex. 2. $\int_0^a \frac{dx}{x^n} = \frac{a^{1-n}}{1-n}$, $R(n) < 1$.

Here a is necessarily real. Let $n = \alpha + i\beta$; then $x^n = x^\alpha e^{i\beta \log x}$ and $|x^n| = x^\alpha$ so that

$$\left|\int_h^k \frac{dx}{x^n}\right| \leqq \int_h^k \frac{dx}{x^\alpha} = \frac{k^{1-\alpha} - h^{1-\alpha}}{1-\alpha}$$

and therefore, if $\alpha = R(n) < 1$, this integral tends to zero when h and k tend to zero. The result then follows since the integral of x^{-n} is $x^{1-n}/(1-n)$.

Ex. 3. $\int_a^b \frac{dx}{x+c} = \log\frac{b+c}{a+c}$, c complex.

Of course a and b are real. The value of $\log\{(b+c)/(a+c)\}$ is the *principal value*, and its amplitude θ is such that $-\pi < \theta \leqq \pi$ (§ 70).

Ex. 4. $\int_0^\infty e^{-cx}x^{n-1}\,dx = \dfrac{\Gamma(n)}{c^n}$, $R(c) > 0$, n real and positive.

Let $c = a + ib$, $a > 0$; by definition

$$\int_0^\infty e^{-cx}x^{n-1}\,dx = \int_0^\infty e^{-ax}x^{n-1}\cos bx\,dx - i\int_0^\infty e^{-ax}x^{n-1}\sin bx\,dx,$$

and therefore (*E.T.* p. 471, equations (6), (7)) the integral is equal to

$$\frac{\Gamma(n)(\cos n\theta - i\sin n\theta)}{r^n} = \Gamma(n)\left(\frac{\cos\theta - i\sin\theta}{r}\right)^n = \frac{\Gamma(n)}{c^n}$$

since $a = r\cos\theta$, $b = r\sin\theta$, $(\cos\theta - i\sin\theta)/r = (a - ib)/r^2 = 1/c$.

The integral also converges if $a = R(c) = 0$ and $0 < n < 1$ (*E.T.* p. 471, (8)),

so that
$$\int_0^\infty e^{-ibx}x^{n-1}\,dx = \frac{\Gamma(n)}{(ib)^n},\quad 0 < n < 1,\ b \text{ real},$$

and the amplitude of ib is $\frac{\pi}{2}$ or $-\frac{\pi}{2}$ according as $b > 0$ or $b < 0$.

Definition of $\Gamma(x)$ *as an integral when* x *is complex,* $R(x) > 0$.

The definition of $\Gamma(x)$ as the limit of the product $P_n(x)$ (§ 95, (1*a*)), where

$$P_n(x) = \frac{n!\,n^x}{x(x+1)(x+2)\dots(x+n)},$$

will now be applied to show that the usual expression for $\Gamma(x)$ as an integral holds for complex values of x if $R(x) > 0$.

Ex. 5. $\int_0^\infty e^{-t}t^{x-1}\,dt = \mathop{\mathcal{L}} P_n(x) = \Gamma(x)$, $R(x) > 0$.

Let $x = \xi + i\eta$, $\xi > 0$; then $|e^{-t}t^{x-1}| = e^{-t}t^{\xi-1}$ so that the integral converges at ∞; by Ex. 2 it also converges at 0 if $\xi > 0$.

Now, if $R(x) = \xi > 0$ and n a positive integer, the integral

$$\int_0^1 s^{x-1}(1-s)^n\,ds$$

is convergent; the value of the integral is easily found to be

$$(n!)/\{x(x+1)(x+2)\dots(x+n)\},$$

and therefore, if $s = t/n$,

$$\int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1}\,dt = P_n(x).$$

Hence
$$\mathop{\mathcal{L}}_{n\to\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1}\,dt = \mathop{\mathcal{L}}_{n\to\infty} P_n(x) = \Gamma(x)$$

It has to be proved that

$$\mathop{\mathcal{L}}_{n\to\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1}\,dt = \int_0^\infty e^{-t}\,t^{x-1}\,dt.$$

Now

$$\int_0^\infty e^{-t}\,t^{x-1}\,dt - \int_0^n \left(1-\frac{t}{n}\right)^n t^{x-1}\,dt$$

$$= \int_0^n \left\{e^{-t} - \left(1-\frac{t}{n}\right)^n\right\} t^{x-1}\,dt + \int_n^\infty e^{-t}\,t^{x-1}\,dt,$$

and it is proved below that if $0 \leqq t \leqq n$,

$$0 \leqq e^{-t} - \left(1-\frac{t}{n}\right)^n \leqq e^{-t}.\frac{t^2}{n},$$

so that if $R(x) = \xi > 0$,

$$0 \leqq \int_0^n \left\{e^{-t} - \left(1-\frac{t}{n}\right)^n\right\} t^{\xi-1}\,dt \leqq \frac{1}{n}\int_0^n e^{-t}\,t^{\xi+1}\,dt < \frac{1}{n}\Gamma(\xi+2).$$

Hence $\displaystyle\int_0^n \left\{e^{-t} - \left(1-\frac{t}{n}\right)^n\right\} t^{\xi-1}\,dt \to 0$ when $n \to \infty$,

and therefore $\displaystyle\int_0^n \left\{e^{-t} - \left(1-\frac{t}{n}\right)^n\right\} t^{x-1}\,dt \to 0$ when $n \to \infty$,

since $|t^{x-1}| = t^{\xi-1}$.

Further, $\displaystyle\int_n^\infty e^{-t}\,t^{x-1}\,dt \to 0$ when $n \to \infty$

since the integral is convergent. Thus

$$\int_0^\infty e^{-t}\,t^{x-1}\,dt - \mathop{\mathcal{L}}_{n\to\infty} \int_0^n \left(1-\frac{t}{n}\right)^n t^{x-1}\,dt = 0,$$

and $\displaystyle\int_0^\infty e^{-t}\,t^{x-1}\,dt = \Gamma(x),\ R(x) > 0.$

To prove the inequalities used above we have by § 25, Ex. 3 (ii),

$$e^{-\frac{t}{n}} > 1 - \frac{t}{n},\ e^{-t} > \left(1-\frac{t}{n}\right)^n,\ \ 0 < e^{-t} - \left(1-\frac{t}{n}\right)^n \text{ if } 0 < t < n.$$

Also by § 25, Ex. 3 (i), $e^t > (1+t/n)^n$ if $0 < t < n$, so that

$$e^{-t} - \left(1-\frac{t}{n}\right)^n = e^{-t}\left[1 - e^t\left(1-\frac{t}{n}\right)^n\right] < e^{-t}\left[1 - \left(1-\frac{t^2}{n^2}\right)^n\right],$$

and therefore by § 11, (2), with $x = 1$ and $y = 1 - t^2/n^2 > 0$,

$$e^{-t} - \left(1-\frac{t}{n}\right)^n < e^{-t}.\,n\left[1 - \left(1-\frac{t^2}{n^2}\right)\right] = e^{-t}\frac{t^2}{n}$$

so that $\displaystyle 0 < e^{-t} - \left(1-\frac{t}{n}\right)^n < e^{-t}.\frac{t^2}{n}$ if $0 < t < n$.

When $t = 0$, it is obvious that the inequalities become equalities.

For this method of proving these inequalities see Whittaker and Watson, *Modern Analysis* (2nd Ed.), p. 236.

148. Miscellaneous Examples. The following method of extending the range of a constant that occurs in an integral may be noted.

Ex. 1. Evaluate $\int_0^\infty \frac{x^{2m}dx}{x^{2n}+1}$ where m and n are positive integers, $m<n$, and deduce that

$$\int_0^\infty \frac{x^{p-1}dx}{1+x}=\frac{\pi}{\sin p\pi},\ 0<p<1.$$

Here
$$\int_0^\infty \frac{x^{2m}dx}{x^{2n}+1}=\tfrac{1}{2}\int_{-\infty}^\infty \frac{x^{2m}dx}{x^{2n}+1},$$

and, since the last integral converges absolutely at ∞ and at $-\infty$, we may take $\xi'=\xi$ (§ 142, (7)) so that

$$\int_{-\infty}^\infty \frac{x^{2m}dx}{x^{2n}+1}=\lim_{\xi\to\infty}\int_{-\xi}^{\xi}\frac{x^{2m}dx}{x^{2n}+1}.$$

Express the integrand as a sum of partial fractions:

$$\frac{x^{2m}}{x^{2n}+1}=\sum_{r=0}^{n-1}\frac{2A_r(x-\cos\theta_r)}{x^2-2x\cos\theta_r+1}+\frac{1}{n}\sum_{r=0}^{n-1}\frac{\sin\theta_r\sin(2m+1)\theta_r}{(x-\cos\theta_r)^2+\sin^2\theta_r},$$

where $\theta_r=(2r+1)\pi/2n$. Now

$$\int_{-\xi}^{\xi}\frac{2(x-\cos\theta_r)dx}{x^2-2x\cos\theta_r+1}=\log\frac{\xi^2-2\xi\cos\theta_r+1}{\xi^2+2\xi\cos\theta_r+1}\to 0 \text{ when } \xi\to\infty,$$

and therefore

$$\int_{-\infty}^\infty\frac{x^{2m}dx}{x^{2n}+1}=\frac{\pi}{n}\sum_{r=0}^{n-1}\sin(2m+1)\theta_r=\frac{\pi}{n}\operatorname{cosec}\frac{(2m+1)\pi}{2n},$$

or
$$\int_0^\infty\frac{x^{2m}dx}{x^{2n}+1}=\frac{\pi}{2n}\operatorname{cosec}\frac{(2m+1)\pi}{2n}.$$

Now change the variable by the substitution $x^{2n}=y$; then

$$\int_0^\infty\frac{y^{\frac{2m+1}{2n}-1}dy}{1+y}=\pi\operatorname{cosec}\frac{(2m+1)\pi}{2n}.$$

If we put $p=(2m+1)/2n$ it is not hard to prove that the integral is a continuous function of p for the range $\delta\leqq p\leqq 1-\eta$, when δ and η are positive and arbitrarily small; since $\operatorname{cosec} p\pi$ is also continuous for the same range, the integral can now be defined to be equal to $\pi\operatorname{cosec} p\pi$ if p is any real number such that $\delta\leqq p\leqq 1-\eta$. (See § 26.)

Ex. 2. If m, n, p are positive integers, $m<p$ and $n<p$, and if $(x^{2m}-x^{2n})/(x^{2p}-1)$ be defined for $x=\pm1$ as the limit of the fraction for x tending to ±1, prove by the method of Ex. 1 that

$$\int_{-\infty}\frac{x^{2m}-x^{2n}}{x^{2p}-1}\,dx=\frac{\pi}{p}\left(\cot\frac{2n+1}{2p}\pi-\cot\frac{2m+1}{2p}\pi\right),$$

and deduce that if $0<a<1$ and $0<b<1$,

$$\int_0^\infty\frac{x^{a-1}-x^{b-1}}{1-x}\,dx=\pi(\cot a\pi-\cot b\pi).$$

The definition of an integral as the limit of a sum is not directly applicable when the integrand is not bounded or when the range of integration is infinite; there are, however, classes of cases in which the limit of a sum may be expressed as an improper integral.

Suppose first that the range (a, b) is finite and that a is the only singular point of $F(x)$ in (a, b); in the notation of § 113 let

$$S_n = \sum_{r=1}^{n} F(a+rh)h, \quad nh = b-a,$$

so that $h \to 0$ when $n \to \infty$. The following theorem may be stated.

THEOREM I. *If $F(x)$ is monotonic and has always the same sign in (a, b) and if the integral of $F(x)$ over (a, b) converges, then*

$$\mathop{\mathcal{L}}_{n\to\infty} S_n = \mathop{\mathcal{L}}_{h\to 0} \sum_{r=1}^{n} F(a+rh)h = \int_a^b F(x)dx.$$

Suppose $F(x)$ to be positive so that $F(x) \to +\infty$ when $x \to a+0$. (If $F(x)$ is negative, let $\varphi(x) = -F(x)$ and the reasoning holds for $\varphi(x)$, and therefore for $F(x)$ by changing sign in each member of the above equation.) Thus $F(x)$ decreases as x increases.

By the monotonic property of $F(x)$ we have

$$\int_{a+(r-1)h}^{a+rh} F(x)dx > F(a+rh)h, \quad r = 1, 2, \ldots, n,$$

and therefore, summing from $r=1$ to $r=n$, since the integral of $F(x)$ converges,

$$\int_a^b F(x)dx > S_n \quad \text{.........................(i)}$$

Again, $\quad F(a+rh)h > \displaystyle\int_{a+rh}^{a+(r+1)h} F(x)dx, \quad r = 1, 2, \ldots (n-1),$

and therefore, summing from $r=1$ to $r=n-1$ and adding the term $F(b)h$ to both sums,

$$S_n > \int_{a+h}^{b} F(x)dx + F(b)h. \quad \text{....................(ii)}$$

Let $n \to \infty$ and $h \to 0$; the inequalities (i) and (ii) give the theorem.

Cor. 1. It is obvious that a similar theorem holds when b is the only singular point of $F(x)$ in (a, b) and $F(x)$ is monotonic and of the same sign in (a, b); in this case

$$S_n = \sum_{r=0}^{n-1} F(a+rh)h.$$

Cor. 2. Further, the monotonic property of $F(x)$ is essential only in an interval (a, c), $a<c<b$, (or in an interval (c, b)); it is not hard to modify the proof to suit this case. Lastly, $F(x)$ might, when $F(x)$ is supposed as in the proof to be positive, be allowed to take negative values in an interval (c, b), $a<c<b$; the theorem would hold for $F(x)+C$, where C is a constant such that $F(x)+C$ is positive, and therefore would hold for $F(x)$.

Ex. 3. $\int_0^{\frac{\pi}{2}} \log\sin x\,dx = -\frac{\pi}{2}\log 2.$

In the identity

$$x^n - 1 = (x-1)\prod_{r=1}^{n-1}\left(x - \cos\frac{2r\pi}{n} - i\sin\frac{2r\pi}{n}\right)$$

after dividing by $x-1$ let $x\to 1$ and take the modulus of each side of the equation; then

$$n = \prod_{r=1}^{n-1}\left(2\sin\frac{r\pi}{n}\right) = \prod_{r=1}^{n-1}\left[2^2\sin\frac{r\pi}{2n}\cdot\sin\frac{(n-r)\pi}{2n}\right],$$

since $\cos(r\pi/2n) = \sin[(n-r)\pi/2n]$. Take the square root (which is positive); therefore

$$\frac{\sqrt{n}}{2^{n-1}} = \sin\frac{\pi}{2n}\cdot\sin\frac{2\pi}{2n}\cdot\sin\frac{3\pi}{2n}\cdot\ldots\sin\frac{(n-1)\pi}{2n}.$$

Let $h=\pi/2n$, take the logarithm of each member and multiply by h; thus

$$S_n = \sum_{r=1}^{n-1} h\log(\sin rh) = \frac{\pi}{4}\frac{\log n}{n} - \frac{\pi}{2}\left(1-\frac{1}{n}\right)\log 2,$$

and $S_n \to -\frac{\pi}{2}\log 2$ when $h\to 0$ or $n\to\infty$. The integral is convergent and its value is therefore $-\frac{\pi}{2}\log 2$.

Cor. $\int_0^1 \log(\sin\pi x)\,dx = \frac{1}{\pi}\int_0^{\pi}\log(\sin x)\,dx = \frac{2}{\pi}\int_0^{\frac{\pi}{2}}\log(\sin x)\,dx = -\log 2.$

Ex. 4. If $0<\varrho<1$ and $\varrho^n\to 0$ when $\varrho\to 1$, prove that

$$\int_0^b \log x\,dx = \underset{\rho\to 1}{\mathcal{L}}\, b(1-\varrho)\sum_{r=0}^{n-1}\varrho^r\log(b\varrho^r) = b\log b - b.$$

Compare *Exercises* XIII, 2, (iii).

THEOREM II. *If $F(x)$ is monotonic in (a,∞) and tends to zero when x tends to infinity; if, further, a is the only singular point of $F(x)$, then*

$$\underset{h\to 0}{\mathcal{L}}\, h\sum_{n=1}^{\infty} F(a+nh) = \int_a^{\infty} F(x)\,dx,$$

provided that the integral converges.

Since $F(x)$ is monotonic and converges to zero it cannot change sign; $F(x)$ will be supposed to be positive. Now let

$$S_n = \sum_{r=1}^{n} F(a+rh)h\,;$$

then, h being kept constant, we have by equation (i) above and the monotonic property of $F(x)$,

$$S_n < \int_a^{a+nh} F(x)\,dx < \int_a^{\infty} F(x)\,dx = K \text{ say,}$$

the integral being convergent and equal to K.

Thus S_n increases as n increases, but is less than the fixed number K; therefore the series S, where

$$S = h\sum_{n=1}^{\infty} F(a+nh),$$

is convergent. It is now easily seen, by the same method as used for the proof of Theorem I, that

$$\int_a^{\infty} F(x)\,dx > h\sum_{n=1}^{\infty} F(a+nh) > \int_{a+h}^{\infty} F(x)\,dx.$$

Let $h \to 0$ and the theorem follows at once.

Ex. 5. Prove that $\int_0^{\infty} x^{\mu-1}e^{-x}\,dx = \Gamma(\mu)$, $\mu > 0$.

Let $S = h\sum_{n=1}^{\infty}(nh)^{\mu-1}e^{-nh} = h^{\mu}\sum_{n=1}^{\infty} n^{\mu-1}e^{-nh}$.

Now let $e^{-h} = y$ so that $y \to 1$ when $h \to 0$ and $h/(1-e^{-h}) \to 1$ when $y \to 1$. The series S may be put in the form

$$S = \left(\frac{h}{1-e^{-h}}\right)^{\mu}(1-y)^{\mu}\{1^{\mu-1}y + 2^{\mu-1}y^2 + \ldots + n^{\mu-1}y^n + \ldots\}$$

and, by § 98, Ex. 5, $S \to \Gamma(\mu)$ when $y \to 1$. The integral is therefore equal to $\Gamma(\mu)$, for the integral is known to converge.

The following test for the convergence of a series Σa_n of positive terms was given by Maclaurin (*Fluxions*, § 350) though it is often called Cauchy's Integral Test.

Integral Test for Convergence of Series. If $a_n \geqq a_{n+1} > 0$ and if $a_n = F(n)$ where $F(x)$ is a positive, monotonic, decreasing function of x, defined for $x \geqq 1$, the series Σa_n and the integral I, where

$$I = \int_1^{\infty} F(x)\,dx,$$

either both converge or both diverge.

A graph of $F(x)$ will make the inequalities used below intuitive.

Let
$$S_n=\sum_{r=1}^{n} a_r,\quad I_n=\int_1^n F(x)\,dx\,;$$
then, from the monotonic property of $F(x)$, it follows at once that
$$a_{r-1} \geqq \int_{r-1}^{r} F(x)\,dx \geqq a_r.$$

Sum from $r=2$ to $r=n$; therefore
$$S_n-a_n \geqq I_n \geqq S_n-a_1,\ \text{or,}\ a_1 \geqq S_n-I_n \geqq a_n>0.$$

Hence (S_n-I_n) is bounded and not negative. Further (S_n-I_n) decreases as n increases; for
$$\begin{aligned}(S_n-I_n)-(S_{n+1}-I_{n+1}) &= (I_{n+1}-I_n)-(S_{n+1}-S_n)\\ &= \int_n^{n+1} F(x)\,dx-a_{n+1} \geqq 0.\end{aligned}$$
Thus (S_n-I_n) tends to a limit which is not negative and does not exceed a_1; therefore S_n and I_n either both tend to a limit or both tend to infinity.

Even if S_n and I_n both tend to infinity the difference (S_n-I_n) tends to a limit.

Ex. 6. The series $\Sigma F(n)$ and $\Sigma e^n F(e^n)$ are either both convergent or both divergent, $F(x)$ being a positive, monotonic, decreasing function of x, defined for $x \geqq 1$. (*Cauchy's Condensation Test*, § 60, Ex. 1.)

Let N be positive and arbitrarily large, $\log b \geqq N$, $c>b$ and $x=e^y$; then
$$\int_b^c F(x)\,dx=\int_{\log b}^{\log c} e^y F(e^y)\,dy$$
and therefore the two integrals are either both convergent or both divergent.

Again, from the monotonic properties of $F(x)$ and e^x,
$$e^{r-1}F(e^r)<\int_{r-1}^{r} e^x F(e^x)\,dx,\quad \frac{1}{e}\sum_{r=n}^{\infty} e^r F(e^r)<\int_{n-1}^{\infty} e^x F(e^x)\,dx,$$
$$e^{r+1}F(e^r)>\int_r^{r+1} e^x F(e^x)\,dx,\quad e\sum_{r=n}^{\infty} e^r F(e^r)>\int_n^{\infty} e^x F(e^x)\,dx.$$
Hence the series $\Sigma e^n F(e^n)$ converges or diverges according as the integral of $e^x F(e^x)$ converges or diverges, therefore according as the integral of $F(x)$ converges or diverges and therefore finally according as the series $\Sigma F(n)$ converges or diverges.

Ex. 7. **Euler's Constant.** If $a_r=1/r$, then $F(x)=1/x$ and
$$S_n-I_n=1+\tfrac{1}{2}+\tfrac{1}{3}+\ldots+\frac{1}{n}-\log n=C_n,\ \text{say.}$$
But (S_n-I_n) tends to a limit, γ say, which lies between 0 and a_1 $(=1)$; therefore $C_n \to \gamma$ when $n \to \infty$. (Exercises II, 8.)

EXERCISES XVII.

Prove that the integrals in Examples 1-6 are convergent.

1. $\int_0^\infty \frac{1}{e^{a/x}-1}\frac{dx}{x^n}$, $a>0$, $n>2$.

2. $\int_0^\infty \left(\frac{1}{x}-\frac{1}{\sinh x}\right)\frac{dx}{x}$.

3. $\int_0^1 x^{m-1}(1-x)^{n-1}\left(\log\frac{1}{x}\right)^p dx$, $m>0$, $n>0$, p a positivo integer.

4. $\int_0^\infty \left(\frac{1}{e^x-1}-\frac{1}{x}+\frac{1}{2}\right)\frac{e^{-kx}}{x}\,dx$, $k>0$.

5. $\int_0^\infty y\,dx$ if $\sinh x \sinh y = 1$.

6. $\int_0^{\frac{\pi}{2}} (\sin x)^{m-1}(\log\sin x)^p dx$, $m>0$, p a positive integer.

7. The threo integrals

$$\int_0^\infty \left(\frac{1}{1+x}-e^{-x}\right)\frac{dx}{x}, \quad \int_0^\infty\left(\frac{1}{1-e^{-x}}-\frac{1}{x}\right)e^{-x}dx, \quad \frac{1}{1+x^2}-e^{-x}\Big)\frac{dx}{x}$$

are convergent and oqual.

8. $\int_0^{\frac{\pi}{2}}\left(\frac{x}{\sin x}\right)^2 dx = \pi\log 2$.

9. $\int_0^\infty \frac{\sin ax \sin bx}{x}\,dx = \frac{1}{2}\log\frac{a+b}{a-b}$, $a>b>0$.

10. $\int_0^\infty \log\left(\frac{m+ne^{-ax}}{m+ne^{-bx}}\right)\frac{dx}{x} = \log\left(1+\frac{n}{m}\right)\log\frac{b}{a}$,

whoro $a>0$, $b>0$, $m>0$, $m+n>0$.

11. If $a>0$, $b>0$ and n a positivo integer,

(i) $\int_0^\infty \frac{\cos^{2n-}\text{[illegible]}}{x}\,dx = \log\frac{b}{a}$;

(ii) $\int_0^\infty \frac{\cos^{2n}ax-\cos^{2n}bx}{x}\,dx = \left(1-\frac{(2n)!}{n!n!}\frac{1}{2^{2n}}\right)\log\frac{b}{a}$.

12. (i) $\int_0^\infty \frac{\cos ax - \cos bx}{x^2}\,dx = \frac{\pi}{2}(b-a)$, $a>0$, $b>0$;

(ii) $\int_0^\infty \frac{\sin ax \sin bx}{x^2}\,dx = \frac{\pi}{2}b$, $a>b>0$.

13. If $\varphi(x)$ and $\varphi'(x)$ satisfy the conditions of Ex. 28, § 146, and if $a>0$, $b>0$, show that

$$\int_0^\infty \left\{\frac{\varphi(ax)-\varphi(bx)}{x^2}-(a-b)\frac{\varphi'(x)}{x}\right\}dx = \varphi'(0)\{b\log b - a\log a + a - b\}.$$

14. If n is a positive integer.

$$\int_0^\infty \frac{\sin^{2n} x}{x^2}\,dx = \frac{1.3.5.\ldots.(2n-3)}{2.4.6.\ldots.(2n-2)}\frac{\pi}{2},\quad n>1$$
$$=\frac{\pi}{2},\quad n=1.$$

15. If $F(x)$ is a positive, monotonic, decreasing function defined for $x \geqq 1$ and if $e^x F(e^x)/F(x) < K < 1$ when $x \geqq G \geqq 1$, prove that the integral of $F(x)$ converges at ∞ and that therefore the series $\Sigma F(n)$ converges.

$$\left[\int_{e^b}^{e^c} F(x)dx = \int_b^c e^y F(e^y)dy < K\int_b^c F(x)dx,\quad c>b>G,\right.$$

$$(1-K)\int_{e^b}^{e^c} F(x)dx < K\left[\int_b^c F(x)dx - \int_{e^b}^{e^c} F(x)dx\right]$$
$$= K\left[\int_b^{e^b} F(x)dx - \int_c^{e^c} F(x)dx\right]$$

so that
$$\int_{e^b}^{e^c} F(x)dx < \frac{K}{1-K}\int_b^{e^b} F(x)dx, \text{ a constant.}$$

Hence, since $F(x)$ is positive, the integral of $F(x)$ converges at ∞ and therefore the series $\Sigma F(n)$ is convergent.]

16. If $F(x)$ is as in Ex. 15, but $e^x F(e^x)/F(x) \geqq 1$ when $x \geqq G$, show that the integral of $F(x)$ diverges at ∞ and that the series $\Sigma F(n)$ is divergent.

The tests given by Ex. 15 and Ex. 16 are known as **Ermakoff's Tests** for convergence and divergence.

Examples 17-22 are from Pólya and Szegö, *Aufgaben*, I, pp. 40-42.

17. If $\alpha > 0$, prove that

(i) $$\lim_{n\to\infty} \frac{1^{\alpha-1}+2^{\alpha-1}+3^{\alpha-1}+\ldots+n^{\alpha-1}}{n^\alpha} = \frac{1}{\alpha};$$

(ii) $$\lim_{n\to\infty} \frac{1^{\alpha-1}-2^{\alpha-1}+3^{\alpha-1}-\ldots+(-1)^{n-1}n^{\alpha-1}}{n^\alpha} = 0.$$

18. Show that if $0<t<1$, $\displaystyle\lim_{t\to1}(1-t)\sum_{n=1}^\infty \frac{t^n}{1+t^n} = \log 2.$

19. Prove that $\displaystyle\lim_{t\to0} t\left\{1+2\sum_{n=1}^\infty \frac{1}{1+n^2t^2}\right\} = \pi.$

20. Prove that, if $0<t<1$,

$$\lim_{t\to1}\left\{(1-t)\sum_{n=1}^\infty \frac{t^n}{1-t^n} + \log(1-t)\right\} = \int_0^\infty \left(\frac{1}{1-e^{-x}} - \frac{1}{x}\right)e^{-x}dx.$$

21. Prove that

(i) $\int_0^\infty \log(1+x^{-a})\,dx = \mathop{\mathcal{L}}_{h\to 0} h \sum_{n=1}^{\infty} \log[1+(nh)^{-a}],\ a>1;$

(ii) if $f(t)=\prod_{n=1}^{\infty}\left[1+\left(\frac{t}{n}\right)^a\right]$, $\mathop{\mathcal{L}}_{t\to\infty} \frac{1}{t}\log f(t) = \pi \operatorname{cosec}(\pi/a)$.

22. Prove that, if $0 \leqq \varphi \leqq \pi$,

$$\int_0^\infty \log(1-2x^{-2}\cos 2\varphi + x^{-4})\,dx = 2\pi\sin\varphi,$$

by considering the integral as the limit

$$\mathop{\mathcal{L}}_{h\to 0} h \sum_{n=1}^{\infty} \log\left\{1-\frac{2\cos 2\varphi}{n^2h^2}+\frac{1}{n^4h^4}\right\}.$$

[In the infinite product for $\sin x/x$, let $x/\pi = e^{i\phi}/h$ and take the square of the modulus of $\sin x/x$ and of the product; this gives

$$\frac{h^2}{4\pi^2}\left\{e^{\frac{2\pi}{h}\sin\phi}+e^{-\frac{2\pi}{h}\sin\phi}-2\cos\left(\frac{2\pi}{h}\cos\varphi\right)\right\}$$
$$=\prod_{n=1}^{\infty}\left(1-\frac{2\cos 2\varphi}{n^2h^2}+\frac{1}{n^4h^4}\right), \text{ etc.}]$$

CHAPTER XIII

IMPROPER INTEGRALS: REPEATED AND DOUBLE INTEGRALS, FIELD OF INTEGRATION FINITE

149. Improper Double Integrals. When the integrand is not bounded or the field of integration not finite, the method employed for defining the improper integral of a function of one variable is again used. The double integral of a bounded function over a finite area is now called, when distinction is necessary, a **proper double integral**; when the definition of the double integral is extended so as to provide for cases in which the integrand is not bounded or the field of integration not finite the integral is called an **improper double integral.**

Closely connected, yet not identical, with the problem of the double integral is that of the repeated integral and, in general, the problem of integrating and differentiating under the sign of integration; when the conditions of §§ 121, 126 no longer hold, the theorems need further examination.

The subject is one of considerable difficulty; an admirable treatment of it in its most general aspect will be found in Hobson's *Functions of a Real Variable*, but that treatment is based on considerations that are outside the scope of this book. All that will be attempted here will be to give an account of the subject as far as it bears on the most frequently occurring types of integrals.

When the integrand $F(x, y)$ tends to infinity if the point (x, y) tends to one or more isolated points or to any point on one or more curves within or on the boundary of the field of integration A, these points and curves are called *points and curves of infinite discontinuity.* Let these points and curves be excluded from the area A by drawing appropriate curves to mark off the several points and curves of discontinuity from the rest of the area A: for example, the lines abc and def in

Fig. 6 of § 124 where EF and GFH are lines of discontinuity, or, the rectangle of sides ε, ε' (*E.T.* p. 447, Fig. 88), or the lines in Fig. 89 (*E.T.* p. 448).

It will be always assumed that the number of isolated points and curves of infinite discontinuity is finite.

Let A' denote the area left in A when the areas enclosed by the *auxiliary curves*, that is, the curves drawn to enclose the points and lines of discontinuity, have been excluded; we then have the definition:

Definition. If the double integral of $F(x, y)$ over A' tends to a limit when A' tends to A, that limit is defined to be the improper double integral of $F(x, y)$ over A.

In the same way, if the double integral of $F(x, y)$ over a finite area tends to a limit when that area tends to infinity—that is, to include a part (or even the whole) of the co-ordinate plane that is unlimited in extent—that limit is defined to be the improper double integral over that infinite area. For an example of a double integral over an infinite area see Ex. 3, p. 340, of the *Elementary Treatise.*

In the cases we discuss the auxiliary curves are of a very simple kind, frequently straight lines; the areas included by these auxiliary curves are usually specified in such a way that they tend to zero when certain numbers, δ, δ', η, ... tend to zero.

The following statement of a simple but typical problem may help the student to appreciate the nature of the difficulties and the particular steps in a proof that require attention.

Problem. Let $F(x, y)=x^{m-1}y^{n-1}f(x, y)$ where m and n are each less than unity and $f(x, y)$ is a continuous function of x and y, and suppose the field of integration to be the triangle T bounded by the lines

$$x=0,\ y=0,\ x+y=c>0 \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(T)$$

The function $F(x, y)$ is not bounded in T and therefore has no proper double integral over T; let the field be contracted to the triangle T' bounded by the lines

$$x=\delta,\ y=\delta',\ x+y=c \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(T')$$

where δ and δ' are arbitrarily small positive numbers, and $F(x, y)$ will have a proper double integral over T'. This double integral can be expressed (§ 126) as a repeated integral, namely,

$$\iint_{T'} F(x, y)\,dx\,dy=\int_{\delta}^{c-\delta'}dx\int_{\delta'}^{c-x}F(x, y)\,dy. \quad\ldots\ldots\ldots\ldots(1)$$

It may be noted that there is no loss of generality in taking the triangle T' as the contracted area; for, whatever be the curve or curves

drawn in the triangle T so as to exclude the two sides $x=0$, $y=0$, the curve or curves so drawn must, if the contracted area tends to T, at some stage lie wholly between the triangle T' and the triangle T; and *vice versa*. A similar remark may be made in other cases.

For brevity, let D denote the double integral and R the repeated integral in equation (1). The questions now arise: when δ and δ' tend (independently) to zero, (i) does D tend to a limit? (ii) does R tend to the repeated integral

$$\int_0^c dx\int_0^{c-x}F(x, y)dy\,? \qquad (2)$$

and (iii) if D tends to a limit, is that limit equal to the integral (2)?

Suppose it to have been proved that D tends to a limit, l say; it does not follow from that fact alone that R tends to the integral (2)—that is, that l is equal to the integral (2). In calculating R the numbers δ and δ' are kept till *after* the integrations with respect to y and x have been effected and only then are they made to tend to zero. On the other hand in the integral (2) δ' and δ have been made to tend to zero *before* the integrations with respect to y and x, and proof is needed that this change of order makes no difference in the result (compare *E.T.* p. 472, foot of page).

An important, and often difficult, step in the proof that the limit of R when δ and δ' tend to zero is the integral (2) is to show that the limit of R is the same as the limit of the integral

$$\int_\delta^{c-\delta'} dx\int_0^{c-x}F(x, y)dy\,; \qquad (3)$$

that is, that the *limit* of R is not altered if in the integral with respect to y the number δ' is made zero *before* the integration with respect to x.

It has next to be proved that when δ and δ' tend to zero the integral (3) tends to the integral (2); if the integral (2) is a definite number it follows that the limit of the double integral D is equal to that number. Even when it can be proved independently of the consideration of R that the limit of D exists it must always be proved, if the improper double integral is to be evaluated by the repeated integral (2), that that repeated integral is the limit of R. In practice the existence of the improper double integral is usually established by showing that R tends to the integral (2).

Again, it is often important to know, quite apart from the question of the double integral, whether the two repeated integrals of $F(x\ y)$ over the area T are equal—that is, whether a change in the order of integration makes no difference in the value of the repeated integral. The other repeated integral over T', which may be called R', is

$$\int_{\delta'}^{c-\delta} dy\int_\delta^{c-y}F(x, y)dx$$

and in this case a step in the proof is to show that the limit is not changed by taking the lower limit of the integral with respect to x as zero, instead of δ, before integrating with respect to y.

150. Absolute Convergence. The auxiliary curves that separate the points and lines of discontinuity from the rest of the area A are only restricted by the condition that the area enclosed by a curve which surrounds an *isolated point* must tend to zero *in all its dimensions* and that the area enclosed by a curve drawn round a *line* of discontinuity must tend to zero.

One consequence of this method of defining the integral is that the improper integral of $F(x, y)$ cannot exist—that is, the limit of the proper double integral cannot be a definite number—unless the proper double integral of $|F(x, y)|$ tends to a definite number. In other words, the improper double integral of $F(x, y)$ will not exist unless that of $|F(x, y)|$ exists, so that improper double integrals are always *absolutely* convergent. It is possible so to define the improper double integral that it need not be absolutely convergent, but the properties of such integrals would be much more restricted than when the integral is defined as in § 149. For a discussion of the whole matter and a proof of the assertions just made the student is referred to Hobson's treatise. The subject is merely mentioned to explain a statement that might puzzle the student, since no such restriction applies to simple integrals.

The following simple example may be of interest. Let O be the origin of co-ordinates, A the point (a, a), where $a > 0$, on the line $y = x$, B the projection of A on the x-axis, C the point $(ma, 0)$, $m > 1$, and D the point (ma, a). If $F(x, y) = \cos y/x$, the double integrals of $F(x, y)$ over the triangle OBA and the quadrilateral $OCDA$ exist and

$$\iint_{OBA} F(x, y)\,dx\,dy = \int_0^a \frac{\sin x}{x}\,dx, \quad \iint_{OCDA} F(x, y)\,dx = \int_0^a \frac{\sin x}{x}\,dx + \sin a \log m.$$

When $a \to \infty$ the first integral tends to $\pi/2$ but the second does not tend to a limit so that the double integral over the infinite sector of angle AOB does not exist. It is easy to verify that the integral of $|F(x, y)|$ over the triangle and over the quadrilateral is divergent.

151. Uniform Convergence. The discussion now to be given of repeated integrals is, like that of Chapter XXI of the *Elementary Treatise*, based on the work of De la Vallée Poussin; see, besides the memoir quoted in the *Elementary Treatise*, his article in the *Journal de Mathématiques* (4th Series), vol. viii, year 1892.

When the limits of the repeated integrals are constant the

area associated with the integrals will be taken, in general definitions and theorems, to be the rectangle R bounded by the lines

$$x=a,\ x=b,\ y=a',\ y=b'. \quad \ldots\ldots\ldots\ldots\ldots\ldots(R)$$

When the limits are not all constant the associated area will be taken to be that represented by Fig. 4 and will usually be referred to as " the area A " ; if any given area does not satisfy the conditions to which the curve $EFGH$ of Fig. 4 is restricted it will be understood that it can be divided into a finite number of areas of the type A.

As regards the integrand $F(x, y)$ it will be assumed to be a continuous function of both variables when the point (x, y) is not a point of infinite discontinuity for the function.

Discontinuities. The number of isolated points and the number of lines of discontinuity will always be supposed to be finite, and further, when a curve of discontinuity is not a straight line parallel to a co-ordinate axis, the assumption will be made that it cannot be met by a line parallel to either axis in more than a finite number of points. As in previous work, it will generally be sufficient to consider *one* line of discontinuity when a theorem is being proved (see the *Note*, § 120).

When the area is the rectangle R the following notations will be used :

$$f(y)=\int_a^b F(x,y)dx,\quad g(x)=\int_{a'}^{b'} F(x,y)dy,$$

$$U=\int_{a'}^{b'} f(y)dy=\int_{a'}^{b'} dy\int_a^b F(x,y)dx,\quad V=\int_a^b g(x)dx=\int_a^b dx\int_{a'}^{b'} F(x,y)dy$$

and the same notation, with the proper changes, will be used when A is the area.

I. Let the area be R and let the lines of discontinuity be parallel to an axis.

If (x_1, y_1) is a point of discontinuity of $F(x, y)$ the condition that the integral $f(y_1)$ should converge at x_1 is by § 143 that, ε having the usual meaning, there should be positive numbers δ and δ' such that the singular integrals (§ 143)

$$\int_\alpha^\beta F(x,y_1)dx \quad \text{and} \quad \int_{\alpha'}^{\beta'} F(x,y_1)dx$$

where $x_1-\delta \leqq \alpha < \beta < x_1$ and $x_1 < \alpha' < \beta' \leqq x_1+\delta'$, will each be

numerically less than ε. If more than one point or if all points on the line $x=x_1$ are points of discontinuity of $F(x, y)$, the integral $f(y)$ is said to *converge uniformly at* x_1 for the range $a' \leqq y \leqq b'$ when each of the singular integrals on the line $x=x_1$ is numerically less than ε for *all* values of y in the range; that is, given ε, the same value of δ and the same value of δ' will secure that the singular integrals are each numerically less than ε, whatever value y may take in the range $a' \leqq y \leqq b'$.

If there is a finite number of lines of discontinuity parallel to the y-axis, $x=x_1$, $x=x_2$, ..., $x=x_m$, and if the integral $f(y)$ converges uniformly for the range $a' \leqq y \leqq b'$ at each of the points $x_1, x_2, \ldots, x_m$, then $f(y)$ is said to converge uniformly for the range $a' \leqq y \leqq b'$, *or*, to converge uniformly "in R."

Uniform Convergence in General. If $a'<c'<b'$ it may happen that the integral $f(y)$ is only uniformly convergent in the closed intervals $(a', c'-\eta)$ and $(c'+\eta', b')$ where η and η' are arbitrarily small positive numbers, the convergence thus ceasing to be uniform at c'. If there are one or more values such as c' but only a finite number of them, the integral $f(y)$ is said to converge uniformly *in general* in the interval (a', b') or in R. If $c'=a'$ then $\eta=0$, while $\eta'=0$ if $c'=b'$.

Similar definitions hold for the integral $g(x)$.

II. Suppose that the discontinuities of $F(x, y)$ in R lie on curves (including straight lines) that cannot be cut in more than a finite number of points by a parallel to either axis.

For definiteness, suppose that $x=\psi(y)$ represents a line of discontinuity and that $\psi(y)$ is a monotonic function so that a line parallel to the x-axis meets it in not more than one point; this line, therefore, can also be represented by an equation of the form $y=\varphi(x)$. The first form is used in discussing the integral $f(y)$ and the second in discussing the integral $g(x)$.

Draw the auxiliary curves $x=\psi(y)-\delta$, $x=\psi(y)+\delta'$ where δ and δ' are positive constants. If the line $x=x_1$ meets the curve $x=\psi(y)$ at y_1, the integral $f(y_1)$ will converge at x_1 if δ and δ' can be chosen so that each of the singular integrals at x_1 is numerically less than ε. When δ and δ' can be chosen so that the singular integrals at *every* point on the curve $x=\psi(y)$ will each be numerically less than ε then the integral $f(y)$ is said to converge uniformly at points on that curve. If there are

more curves of discontinuity a similar definition applies for each, and if $f(y)$ converges uniformly for each curve it is said to converge uniformly in the rectangle R.

A similar definition holds for the integral $g(x)$; in this case the auxiliary curves will be of the form $y=\varphi(x)-\delta$, $y=\varphi(x)+\delta'$.

III. The field is the area A, bounded by the curve C or $EFGH$ (Fig. 7). In this case

$$f(y)=\int_{x_1}^{x_2} F(x, y)dx, \quad g(x)=\int_{y_1}^{y_2} F(x, y)dy,$$

where (§ 126) $x_1=NR=\psi_1(y)$, $x_2=NS=\psi_2(y)$, $y_1=MP=\varphi_1(x)$ and $y_2=MQ=\varphi_2(x)$.

The preceding definitions are readily extended to include these integrals. The chief point to note is that the limits x_1 and x_2 are not constants but functions of y and the limits y_1 and y_2 are functions of x. Again, if $x_1=\psi_1(y)$ were a curve of discontinuity only one auxiliary line, $x=\psi_1(y)+\delta'$, would be required for that part of the curve.

The various methods that are used for determining the convergence of the integral of $F(x)$ at a point of discontinuity are of course applicable in the cases just stated; the essential point is that the numbers δ and δ' must be such that the singular integrals, at x_1 say, must be each numerically less than ε, whatever value y may have in the range $a' \leqq y \leqq b'$.

Note It is perhaps worth observing that if the integral $f(y)$ converges uniformly for the range $a' \leqq y \leqq b'$ so does the integral $\int_a^c F(x, y)dx$, where $a<c<b$.

Ex. 1. If $f(x, y)$ is a continuous function of x and y and $0<m<1$, $0<n<1$, $0<p<1$, determine the nature of the convergence of $f(y)$ and $g(x)$ in R when

(i) $F(x, y)=(x-a)^{m-1}(b'-y)^{n-1}f(x, y)$;

(ii) $F(x, y)=(x-a)^{m-1}(b-x)^{p-1}(y-a')^{n-1}f(x, y)$.

In all cases there is uniform convergence in general. For (i) $f(y)$ only converges uniformly for $a' \leqq y \leqq \beta' < b'$ and $g(x)$ only for $a<\alpha \leqq x \leqq b$.

For (ii) $f(y)$ ceases to converge uniformly at a' and $g(x)$ at a and b.

Ex. 2. The integrals $\int_0^1 e^{-x}x^{y-1}dx$, $\int_0^1 e^{-x}x^{y-1}(\log x)^m dx$,

where m is a positive integer, converge absolutely and uniformly for the range $0<c \leqq y \leqq d$ where d is any positive number.

See § 146, Ex. 14 and Ex. 17.

152. Continuity of Integrals. When a curve of discontinuity is not a straight line parallel to an axis the assumption will be made that it cannot be cut in more than *one* point by a line parallel to either axis and that it may therefore be represented by either of the equations $x=\psi(y)$ or $y=\varphi(x)$; since, by hypothesis, no curve can be met by a line parallel to either axis in more than a finite number of points, any curve can be divided into a finite number of parts each of which satisfies the above condition. Further, in proving any theorem it will usually be sufficient to prove it for only one curve of discontinuity, and it may be assumed that the curve begins and ends on the boundary of the field ; if it does not begin or end on a boundary, a supplementary auxiliary curve parallel to an axis can be introduced so as to close the area within which the curve lies.

THEOREM I. *If the integrals $f(y)$ and $g(x)$ are uniformly convergent in R or in A they are continuous functions of y and x respectively for the ranges* $a' \leqq y \leqq b'$ and $a \leqq x \leqq b$.

Take the integral $f(y)$ and draw the auxiliary curves $x=\psi(y)-\delta$ and $x=\psi(y)+\delta'$, thus determining a strip S within which the curve of discontinuity lies. Since the integral $f(y)$ converges uniformly, δ and δ' may be so chosen that when (x, y) is *any* point in S the contribution to $f(y)$ from each of the singular integrals will be numerically less than ε. Therefore, if y_1 and y_2 are the ordinates of any two points in S, the contributions to $|f(y_1)|$ and $|f(y_2)|$ will in each case be less than 2ε, and the contribution to $|f(y_1)-f(y_2)|<4\varepsilon$. When δ and δ' have been thus chosen they are to be kept fixed.

Next, in the area outside the strip S the function $F(x, y)$ is continuous, and therefore also the corresponding part of the integral $f(y)$, so that η may be chosen to make the contribution to $|f(y_1)-f(y_2)|<\varepsilon$ if $|y_1-y_2|<\eta$. Hence, if y_1 and y_2 are two points in the closed interval (a', b') it is possible to choose η so that, provided $|y_1-y_2|<\eta$,

$$|f(y_1)-f(y_2)|<4\varepsilon+\varepsilon=5\varepsilon,$$

and therefore $f(y)$ is continuous for $a' \leqq y \leqq b'$.

A similar proof holds for the integral $g(x)$.

The next theorem is important in determining the existence

of the integrals of $f(y)$ and $g(x)$ when these functions are only uniformly convergent in general.

THEOREM II. (i) *If the integral $f(y)$ is only uniformly convergent in general for the range $a' \leqq y \leqq b'$, but if the integral $u(\xi)$, where*

$$u(\xi) = \int_{a'}^{b'} dy \int_a^\xi F(x, y)\,dx,$$

converges uniformly for the range $a \leqq \xi \leqq b$, then $u(\xi)$ is a continuous function of ξ for that range.

(ii) *If the integral $g(x)$ is only uniformly convergent in general for the range $a \leqq x \leqq b$ but if the integral $v(\eta)$, where*

$$v(\eta) = \int_a^b dx \int_{a'}^\eta F(x, y)\,dy,$$

converges uniformly for the range $a' \leqq \eta \leqq b'$, then $v(\eta)$ is a continuous function of η for that range.

Case (ii) is merely stated for convenience of reference as it is simply Case (i) with the *rôles* of x and y interchanged. For the sake of clearness the symbols ξ and η are taken as the upper limits of the integrals with respect to x and y respectively, though in practice it is quite common to put x for ξ and y for η.

The meaning of the theorem is perhaps made more evident if it be observed that $u(\xi)$ is the repeated integral of $F(x, y)$ over that part of the rectangle R between the side $x = a$ and the line $x = \xi$; as ξ varies from a to b the integral $u(\xi)$ will vary continuously from 0 to U.

Take Case (i) and suppose that the integral $f(y)$ ceases to converge uniformly for the one value c' of y, where $a' < c' < b'$; draw the auxiliary lines $y = c' - \delta$ and $y = c' + \delta'$.

Let $\quad f(\xi, y) = \int_a^\xi F(x, y)\,dx, \quad u(\xi) = \int_{a'}^{b'} f(\xi, y)\,dy;$

then the uniform convergence of the integral $u(\xi)$ means that the singular integrals of $u(\xi)$ at c' may each be made, by choice of δ and δ', numerically less than ε for *every* value of ξ in the range $a \leqq \xi \leqq b$. These singular integrals are, numerically,

$$\left|\int_\alpha^\beta f(\xi, y)\,dy\right|, \; c' - \delta \leqq \alpha < \beta < c', \; \left|\int_{\alpha'}^{\beta'} f(\xi, y)\,dy\right|, \; c' < \alpha' < \beta' \leqq c' + \delta'.$$

Let δ and δ' be chosen so that each of these will be less than ε and be then kept fixed.

The integrals

$$\int_{a'}^{c'-\delta} f(\xi, y)dy \text{ and } \int_{c'+\delta'}^{b'} f(\xi, y)dy$$

converge uniformly for the range $a \leqq \xi \leqq b$ and are therefore continuous functions of ξ; it follows, therefore, as in Theorem I, that $u(\xi)$ is continuous for the same range of ξ.

The *value* $u(b)$ of the continuous function $u(\xi)$ is the repeated integral U and, in the same way, $v(b')$ is the repeated integral V.

THEOREM III. *If $F(x, y)$ does not change sign in the rectangle R, so that $F(x, y)$ is either positive or zero or else negative or zero, then,* (i) *provided the repeated integral U exists and the integral $f(y)$ converges uniformly in general in R, the integral $u(\xi)$ of Theorem II converges uniformly for $a \leqq \xi \leqq b$*; (ii) *provided the repeated integral V exists and the integral $g(x)$ converges uniformly in general in R, the integral $v(\eta)$ of Theorem II converges uniformly for $a' \leqq \eta \leqq b'$.*

Suppose, first, that $F(x, y)$ is never negative in R and take the integral $u(\xi)$; then $u(\xi)$ is a monotonic, increasing function of ξ and must therefore either tend to a limit or tend to infinity. But, by hypothesis, $u(b) = U$, a definite number, and therefore $u(\xi)$ tends to a limit when $\xi \to b$. Further $u(\xi)$ must tend to a limit when $\xi \to \xi_1$ where $a \leqq \xi_1 < b$; because, as before, $u(\xi)$ must either tend to a limit or tend to infinity, and it cannot tend to infinity since $u(\xi_1) < u(b)$ if $\xi_1 < b$ and therefore $u(\xi_1) < U$. Hence $u(\xi)$ is a well-defined function for the range $a \leqq \xi \leqq b$.

With the notation of Theorem II we now have, for the singular integrals at c'

$$\int_{a}^{\beta} f(\xi, y)dy \leqq \int_{a}^{\beta} f(b, y)dy, \quad \int_{a'}^{\beta'} f(\xi, y)dy \leqq \int_{a'}^{\beta'} f(b, y)dy$$

since $f(\xi, y) \leqq f(b, y)$ when $F(x, y)$ is not negative. But $f(b, y)$ does not depend on ξ and therefore, since the integral $u(b)$ converges, the singular integrals satisfy the conditions of Theorem II so that $u(\xi)$ converges uniformly for $a \leqq \xi \leqq b$ and is a continuous function of ξ for that range.

In a similar way the theorem is proved when $F(x, y)$ is not positive; or it follows from the above proof for the function $F_1(x, y)$ where $F_1(x, y) = -F(x, y)$ and therefore for $F(x, y)$. The theorem (ii) is proved in the same way.

Ex. If $F(x, y) = (x-a)^{m-1}(b'-y)^{n-1}f(x, y)$, where $0 < m < 1$, $0 < n < 1$, and $f(x, y)$ is continuous in R, show that $u(\xi)$ and $v(\eta)$ are continuous functions.

153. Change of Order of Integration. In this article the function $F(x, y)$ is assumed to be continuous *except on a finite number of straight lines parallel to the axes*; further, the limits of all the integrals are *constants* so that the field of integration is the rectangle R. The case of constant limits when the above restriction on the lines of discontinuity is removed is considered in § 154.

The notation is

$$f(y) = \int_a^b F(x, y)\,dx, \quad g(x) = \int_{a'}^{b'} F(x, y)\,dy,$$

$$U = \int_{a'}^{b'} f(y)\,dy = \int_{a'}^{b'} dy \int_a^b F(x, y)\,dx,$$

$$V = \int_a^b g(x)\,dx = \int_a^b dx \int_{a'}^{b'} F(x, y)\,dy.$$

THEOREM I. *If all the lines of discontinuity in R are parallel to only one axis then the repeated integrals U and V exist and are equal either* (i) *if the integral $f(y)$ converges uniformly in R when the lines of discontinuity are parallel to the y-axis*, or (ii) *if the integral $g(x)$ converges uniformly when the lines of discontinuity are parallel to the x-axis.*

Take the integral $f(y)$ and suppose that there is only one line of discontinuity parallel to the y-axis, say $x = a$. By § 152, Theorem I, $f(y)$ is a continuous function of y for the range $a' \leqq y \leqq b'$, and therefore the repeated integral U exists.

Now draw the auxiliary line $x = a + \delta$. In the part of the rectangle R that lies between the lines $x = a + \delta$ and $x = b$ the function $F(x, y)$ is continuous and therefore

$$\int_{a+\delta}^b dx \int_{a'}^{b'} F(x, y)\,dy = \int_{a'}^{b'} dy \int_{a+\delta}^b F(x, y)\,dx$$

$$= U - \int_{a'}^{b'} dy \int_a^{a+\delta} F(x, y)\,dx \quad \text{.........(1)}$$

But $f(y)$ converges uniformly in R and therefore δ can be chosen so that if $a' \leqq y \leqq b'$

$$\left| \int_a^{a+\delta} F(x, y)\,dx \right| < \varepsilon, \quad \left| \int_{a'}^{b'} dy \int_a^{a+\delta} F(x, y)\,dx \right| < (b' - a')\,\varepsilon.$$

Hence

$$\mathop{\mathcal{L}}_{\delta\to 0}\int_{a'}^{b'} dy \int_{a}^{a+\delta} F(x, y)dx = 0, \quad \mathop{\mathcal{L}}_{\delta\to 0}\int_{a'}^{b'} dy \int_{a+\delta}^{b} F(x, y)dx = U,$$

and therefore

$$\mathop{\mathcal{L}}_{\to 0}\int_{a+\delta}^{b} dx \int_{a'}^{b'} F(x, y)dy = U;$$

that is,

$$V = U.$$

The proof is obviously such that it may be extended to the case of any finite number of lines parallel to the axis of y; the proof when the lines are all parallel to the x-axis is carried out in the same way.

Ex. 1. $F(x, y) = (x-a)^{m-1}(b-x)^{n-1}f(x, y)$
or $F(x, y) = (y-a')^{m-1}(b'-y)^{n-1}f(x, y)$
where $0 < m < 1$, $0 < n < 1$ and $f(x, y)$ is continuous in R.

Carry out the proof fully by dividing R into two rectangles, in the first case by a line $x = c$, $a < c < b$, and in the second by a line $y = c'$, $a' < c' < b'$; *or*, as one theorem by two auxiliary lines, $x = a + \delta$, $x = b - \delta'$ in the first case and $y = a' + \delta$, $y = b' - \delta'$ in the second.

Notation. For brevity, the rectangle bounded by the lines $x = a_1$, $y = b_1$ and $x = a_2$, $y = b_2$ will sometimes be denoted by $(a_1, b_1; a_2, b_2)$. (§ 119, at end.)

THEOREM II. *If one of the integrals $f(y)$ and $g(x)$ is only uniformly convergent in general in R but the other uniformly convergent in R, then the repeated integrals U and V exist, and $U = V$.*

Let $x = a$ and $y = a'$ be the only lines of discontinuity of $F(x, y)$ in R and suppose that $f(y)$ converges uniformly for the range $a' \leqq y \leqq b'$ while $g(x)$ only converges uniformly for $a < a + \delta \leqq x \leqq b$ where δ is arbitrarily small.

In the rectangle $(a + \delta, a'; b, b')$ the only discontinuities of $F(x, y)$ lie on the line $y = a'$ and the integral $g(x)$ converges uniformly in this rectangle; therefore, by Theorem I,

$$\int_{a+\delta}^{b} dx \int_{a'}^{b'} F(x, y)dy = \int_{a'}^{b'} dy \int_{a+\delta}^{b} F(x, y)dx.$$

But the integral $f(y)$ converges uniformly for $a' \leqq y \leqq b'$ and therefore, as in Theorem I,

$$\mathop{\mathcal{L}}_{\delta\to 0}\int_{a+\delta}^{b} dx \int_{a'}^{b'} F(x, y)dy = \mathop{\mathcal{L}}_{\delta\to 0}\int_{a'}^{b'} dy \int_{a+\delta}^{b} F(x\, y)dx = U, \quad \text{....(2)}$$

that is,

$$V = U.$$

Ex. 2. Carry out the proof in detail (i) when $g(x)$ ceases to converge uniformly for $x=a$, $x=c$ and $x=b$ while $y=a'$, $y=c'$, $y=b'$ are lines of discontinuity of $F(x, y)$; (ii) when $f(y)$ ceases to converge uniformly for $y=a'$, $y=c'$, $y=b'$ while $g(x)$ converges uniformly and $x=a$, $x=c$, $x=b$ are lines of discontinuity of $F(x, y)$.

THEOREM III. *If each of the integrals $f(y)$ and $g(x)$ is only uniformly convergent in general in R but if* one *of the integrals $u(\xi)$ and $v(\eta)$ where*

$$u(\xi)=\int_{a'}^{b'}dy\int_a^{\xi}F(x, y)dx,\quad v(\eta)=\int_a^b dx\int_{a'}^{\eta}F(x, y)dy$$

converges uniformly, $u(\xi)$ for the range $a \leqq \xi \leqq b$ and $v(\eta)$ for the range $a' \leqq \eta \leqq b'$, then the repeated integrals U and V exist and $U=V$.

Let $u(\xi)$ converge uniformly and suppose that the integral $g(x)$ ceases to converge uniformly for the one value $x=c$, where $a<c<b$; then in the rectangles $(a, a';\ \ c-\delta, b')$ and $(c+\delta', a';\ b, b')$, where δ and δ' have the usual meaning, the integral $g(x)$ converges uniformly and therefore, by Theorem II,

$$\int_a^{c-\delta}dx\int_{a'}^{b'}F(x, y)dy+\int_{c+\delta'}^b dx\int_{a'}^{b'}F(x, y)dy$$
$$=\int_{a'}^{b'}dy\int_a^{c-\delta}F(x, y)dx+\int_{a'}^{b'}dy\int_{c+\delta'}^b F(x, y)dx \quad\text{.........(3)}$$

The two repeated integrals in the right-hand member of this equation may be expressed in the form

$$u(c-\delta)+[u(b)-u(c+\delta')]=U+u(c-\delta)-u(c+\delta'),$$

and therefore, since $u(\xi)$ is, by § 152, Theorem II, a continuous function of ξ, the limit of this expression when δ and δ' tend to zero is U. Hence

$$\underset{\delta\to 0}{\mathcal{L}}\int_a^{c-\delta}dx\int_{a'}^{b'}F(x, y)dy+\underset{\delta'\to 0}{\mathcal{L}}\int_{c+\delta'}^b dx\int_{a'}^{b'}F(x, y)dy=U,$$

that is, $$V=U.$$

The student should note the reasons for the statement that the limit for δ and δ' tending to zero of the integrals in the right-hand member of equation (3) is obtained by simply making δ and δ' zero in the integrals. He should also note that in the case of equations (1) and (2) in Theorems I and II it is essentially the continuity of the integral $f(y)$ (due to its

uniform convergence) that makes the *limit* for δ tending to zero of the repeated integral in the right-hand member of these equations equal to the *value* of the integral for δ equal to zero.

Cor. A very important particular case of Theorem III is that for which the function $F(x, y)$ does not change sign in R, because the functions $u(\xi)$ and $v(\eta)$ are, by § 152, Theorem III, continuous functions of ξ and η respectively when the repeated integrals U and V exist.

Ex. 3. If $F(x, y) = x^{m-1}y^{n-1}(1-x)^{p-1}(1-y)^{q-1}f(x, y)$, where all the indices m, n, p, q lie between zero and unity and the function $f(x, y)$ is continuous in the square $(0, 0; 1, 1)$, show that

$$\int_0^1 dy \int_0^1 F(x, y)dx = \int_0^1 dx \int_0^1 F(x, y)dy. \quad \text{.................(i)}$$

Both $f(y)$ and $g(x)$ converge uniformly in general; the integral

$$\int_0^1 x^{m-1}(1-x)^{p-1}f(x, y)dx$$

is a continuous function of y and therefore the integral

$$\int_0^1 y^{n-1}(1-y)^{q-1}dy \int_0^1 x^{m-1}(1-x)^{p-1}f(x, y)dx$$

converges. If $f(x, y)$ did not change sign, the corollary of Theorem III would secure the validity of equation (i); in any case it is easy to prove that $u(\xi)$ is continuous.

154. Change of Order: Variable Limits. It will now be supposed that the field of integration is the area A, bounded by the curve C or $EFGH$ (Fig. 7); the limits of the integrals are no longer constants, though the conclusions hold when the limits are constants and the area of integration is the rectangle R. The lines of discontinuity are not in this article restricted, as in § 153, to straight lines parallel to the axes, and the theorems now to be proved are therefore extensions of those in the preceding article.

The notation is

$$f(y) = \int_{x_1}^{x_2} F(x, y)dx, \qquad g(x) = \int_{y_1}^{y_2} F(x, y)dy,$$

$$U = \int_{a'}^{b'} dy \int_{x_1}^{x_2} F(x, y)dx, \quad V = \int_a^b dx \int_{y_1}^{y_2} F(x, y)dy,$$

where $x_1 = \psi_1(y)$, $x_2 = \psi_2(y)$ and $y_1 = \varphi_1(x)$, $y_2 = \varphi_2(x)$.

For the rectangle R we put $x_1 = a$ and $y_1 = a'$.

The curves of discontinuity will be understood to satisfy the conditions stated at the beginning of § 152.

THEOREM I. *If the integrals $f(y)$ and $g(x)$ converge uniformly in the area A (or, in the rectangle R), the repeated integrals U and V exist and $U = V$.*

By § 152, Theorem I, $f(y)$ and $g(x)$ are continuous functions of y and x respectively for the ranges $a' \leqq y \leqq b'$ and $a \leqq x \leqq b$, and therefore the integrals U and V exist. It has to be proved that $U = V$.

As in § 152, Theorem I, let there be only one curve of discontinuity and enclose it in the strip S (see the proof of the theorem in question). In the parts of A that lie outside S the function $F(x, y)$ is continuous and therefore, when the point (x, y) lies in these parts, the repeated integrals exist and are equal. Hence if, when the point (x, y) lies in S, the repeated integrals tend to zero when the area of S tends to zero, the repeated integrals will be equal when the point (x, y) lies anywhere in A.

Now it was shown in the proof of Theorem I, § 152, that the contributions to $|f(y)|$ and $|g(x)|$ when the point (x, y) is anywhere in S can each be made less than 2ε by choice of δ and δ', and therefore the contributions to the repeated integrals, when δ and δ' have been chosen, will be less than $(b' - a') \cdot 2\varepsilon$ and $(b - a) \cdot 2\varepsilon$ respectively. Hence the contributions to the repeated integrals when the point (x, y) lies in S tend to zero when the area of S tends to zero and therefore $U = V$.

Ex. 1. If $F(x, y) = (x - y)^{m-1} f(x, y)$, show that

$$\int_0^1 dy \int_y^1 F(x, y)dx = \int_0^1 dx \int_0^x F(x, y)dy$$

where $0 < m < 1$ and $f(x, y)$ is continuous in the field of integration.

Here $|f(x, y)| < K$ and for the singular integrals we have

$$\left|\int_y^{y+\delta} F(x, y)dx\right| < K\frac{\delta^m}{m}, \quad \left|\int_{x-\eta}^{x} F(x, y)dy\right| < K\frac{\eta^m}{m}$$

so that the integrals $f(y)$ and $g(x)$ converge uniformly. The theorem therefore applies in this case—a case of Dirichlet's Formula (§ 130, Ex. 10).

Ex. 2. If $F(x, y) = (c - x - y)^{p-1} f(x, y)$, show that $(c > 0)$

$$\int_0^c dy \int_0^{c-y} F(x, y)dx = \int_0^c dx \int_0^{c-x} F(x, y)dy,$$

where $0 < p < 1$ and $f(x, y)$ is continuous in the field of integration.

Note that, if $|f(x, y)| < K$ and $\delta_1 > \delta_2 > 0$,

$$\left|\int_{c-y-\delta_1}^{c-y-\delta_2} F(x, y)dx\right| < \frac{K}{p}(\delta_1^p - \delta_2^p)$$

and therefore tends to zero when $\delta_1 \to 0$ for every value of y in the interval $(0, c)$, so that the integral $f(y)$ converges uniformly. Similarly for the integral $g(x)$.

Notation. If the line $x = \xi = OM$ (Fig. 7) meets the curve $EFGH$ at P and Q, thus dividing the area A into the two parts —A_1, bounded by the arc $QHEP$ and the straight line PQ, and A_2, bounded by the straight line PQ and the arc $PFGQ$—the repeated integral U will be expressed as the sum of the two integrals

$$\int_{a'}^{b'} dy \int_{x_1}^{\xi} F(x, y)dx \text{ and } \int_{a'}^{b'} dy \int_{\xi}^{x_2} F(x, y)dx, \text{ ....}(\alpha)$$

where $x_1 = \psi_1(y)$ is the equation of the arc GHE and $x_2 = \psi_2(y)$ that of the arc EFG.

When $\xi = b$ the second of the integrals (α) is zero and the first of these integrals will be taken to mean the repeated integral U.

Similarly, if $y = \eta$ represents the line NRS, the repeated integral V will be expressed as the sum of the two integrals

$$\int_{a}^{b} dx \int_{y_1}^{\eta} F(x, y)dy \text{ and } \int_{a}^{b} dx \int_{\eta}^{y_2} F(x, y)dy, \text{ ....}(\beta)$$

where $y_1 = \varphi_1(x)$ is the equation of the arc HEF and $y_2 = \varphi_2(x)$ that of the arc HGF · while if $\eta = b'$ the second of the integrals (β) is zero and the first is taken to mean the repeated integral V.

Theorem II. *If one of the integrals $f(y)$ and $g(x)$ is only uniformly convergent in general in A (or, in the rectangle R), but the other converges uniformly in A, then the repeated integrals* U and V *exist, and* $U = V$.

Suppose that the integral $f(y)$ converges uniformly in A but that the integral $g(x)$ only converges uniformly in the intervals $(a, c-\delta)$ and $(c+\delta', b)$ where δ and δ' are positive and arbitrarily small; then the integral U exists, since $f(y)$ converges uniformly. The integral $g(x)$ converges uniformly for the ranges $a \leqq x \leqq c-\delta$ and $c+\delta' \leqq x \leqq b$, and therefore, by Theorem I,

$$\begin{aligned}&\int_{a}^{c-\delta} dx \int_{y_1}^{y_2} F(x, y)dy + \int_{c+\delta'}^{b} dx \int_{y_1}^{y_2} F(x, y)dy \\ &= \int_{a'}^{b'} dy \int_{x_1}^{c-\delta} F(x, y)dx + \int_{a'}^{b'} dy \int_{c+\delta'}^{x_2} F(x, y)dx. \qquad (1)\end{aligned}$$

so that

$$\mathop{L}_{\delta\to 0}\int_a^{c-\delta} dx\int_{y_1}^{y_2} F(x,y)\,dy+\mathop{L}_{\delta'\to 0}\int_{c+\delta'}^{b} dx\int_{y_1}^{y_2} F(x,y)\,dy=U;$$

that is,

$$\int_a^{c} dx\int_{y_1}^{y_2} F(x,y)\,dy+\int_c^{b} dx\int_{y_1}^{y_2} F(x,y)\,dy=U,$$

or

$$V=U.$$

The student should work out the case in which $v(\eta)$ is given as converging uniformly in (a', b').

Cor. If $F(x, y)$ does not change sign in A, then $u(\xi)$ and $v(\eta)$ are continuous functions of ξ and η respectively when the repeated integrals U and V exist (Theorem III); in this case the application of Theorem IV is simplified, since it is only the convergence of the integrals U and V that requires investigation and not the uniform convergence of $u(\xi)$ and $v(\eta)$.

Ex. 5. If $F(x, y)=x^{m-1}y^{n-1}(c-x-y)^{p-1}f(x, y)$, show that $(c>0)$

$$\int_0^c dy\int_0^{c-y} F(x, y)\,dx=\int_0^c dx\int_0^{c-x} F(x, y)\,dy$$

when the indices m, n, p all lie between zero and unity and $f(x, y)$ is continuous in the field of integration.

The convergence of $f(y)$ and $g(x)$ ceases to be uniform only for $y=0$ and $x=0$; the integral

$$\int_0^{c-y} x^{m-1}(c-x-y)^{p-1}f(x, y)\,dx$$

is a continuous function of y and therefore the integral

$$U=\int_0^c y^{n-1}dy\int_0^{c-y} x^{m-1}(c-x-y)^{p-1}f(x, y)\,dx$$

is convergent; if $f(x, y)$ does not change sign the corollary of Theorem IV applies, while if $f(x, y)$ has not always the same sign it is almost "obvious" that $u(\xi)$ converges uniformly.

155. Differentiation under the Integral Sign. Let $f(y)$ and $\varphi(y)$ denote the integrals

$$f(y)=\int_a^b F(x, y)\,dx,\quad \varphi(y)=\int_a^b \frac{\partial F(x, y)}{\partial y}\,dx,$$

and let R_1 be the rectangle $(a, a'; b_1, b')$, where $b_1 \geqq b > a$.

Suppose that, when a, a', b, b' are fixed constants and R is the rectangle $(a, a'; b, b')$, the following conditions are satisfied:

(i) $F(x, y)$ is a continuous function of x and y in R;

(ii) $\partial F/\partial y$ is a continuous function of x and y in R, except at points on a finite number of curves of discontinuity which satisfy the conditions stated in § 152, and at these points the discontinuities of $\partial F/\partial y$ are infinite;

(iii) The integral $\varphi(y)$ converges uniformly in R.

When these conditions are satisfied $\varphi(y)$ is the derivative of $f(y)$, that is,

$$\frac{df(y)}{dy}=\frac{d}{dy}\int_a^b F(x, y)\,dx=\int_a^b \frac{\partial F(x, y)}{\partial y}\,dx.$$

Let v be any fixed value of y in (a', b') and let $(v+h)$ be also in (a', b'); then

$$\frac{df(v)}{dv}=\mathop{\mathcal{L}}_{h\to 0}\frac{1}{h}\int_a^b \{F(x, v+h)-F(x, v)\}\,dx. \quad\text{............(1)}$$

Now $F(x, v)$ is continuous in R and so is $\partial F(x, y)/\partial y$ except at the points in which a line $y=$constant meets the curves of discontinuity; therefore, by § 145, Theorem D,

$$F(x, v+h)-F(x, v)=\int_v^{v+h}\frac{\partial F(x, y)}{\partial y}\,dy.$$

Hence

$$\int_a^b \{F(x, v+h)-F(x, v)\}\,dx=\int_a^b dx\int_v^{v+h}\frac{\partial F(x, y)}{\partial y}\,dy. \quad\text{......(2)}$$

Again, the integral $\varphi(y)$ converges uniformly in R, and the integral

$$\int_v^{v+h}\frac{\partial F(x, y)}{\partial y}\,dy$$

is a continuous function of x in (a, b); therefore, by § 154, Theorem I,

$$\int_a^b dx\int_v^{v+h}\frac{\partial F(x, y)}{\partial y}\,dy=\int_v^{v+h}dy\int_a^b\frac{\partial F(x, y)}{\partial y}\,dx=\int_v^{v+h}\varphi(y)\,dy, \quad (3)$$

so that, by (1), (2) and (3),

$$\frac{df(v)}{dv}=\mathop{\mathcal{L}}_{h\to 0}\frac{1}{h}\int_v^{v+h}\varphi(y)\,dy=\varphi(v)=\int_a^b\frac{\partial F(x, v)}{\partial v}\,dx,$$

or

$$\frac{d}{dy}\int_a^b F(x, y)\,dx=\frac{df(y)}{dy}=\int_a^b\frac{\partial F(x, y)}{\partial y}\,dx,$$

since v is any value of y in (a', b').

Suppose next that b is not constant but is a function of y, say $b=\psi(y)$. Let $F(x, y)$ be a continuous function of x and y in R_1 (a, a', b_1, b'), and let $\partial F/\partial y$ also be continuous in R_1 except on the curve $x=b=\psi(y)$, the discontinuities of $\partial F/\partial y$

at points on this curve being infinite. The functions b or $\psi(y)$ and db/dy or $\psi'(y)$ are to be continuous for $a' \leqq y \leqq b'$.

If $a \leqq x \leqq \beta < b$, formula (2) of § 121 gives

$$\frac{d}{dy}\int_a^\beta F(x, y)dx = \int_a^\beta \frac{\partial F(x, y)}{\partial y}dx + F(\beta, y)\frac{d\beta}{dy}.$$

If now the condition that the integral $\int_a^b \frac{\partial F(x, y)}{\partial y}dx$ converges uniformly for the range $a' \leqq y \leqq b'$ is satisfied, the integral will be a continuous function of y, and therefore, since $F(x, y)$ and db/dy are continuous, we find by letting β tend to b that

$$\frac{d}{dy}\int_a^b F(x, y)dx = \int_a^b \frac{\partial F(x, y)}{\partial y}dx + F(b, y)\frac{db}{dy},$$

which is the usual formula.

A similar discontinuity for $x = a = \psi_1(y)$ may be treated in the same way, the rectangle R_1 being $(a_1, a'; b, b')$, where $a \leqq a = \psi_1(y)$.

A change of the variable of integration is often effective in simplifying the problem of differentiating under the integral sign; see Ex. 2 below.

Ex. 1. If $f(y) = \int_0^{\frac{\pi}{2}} (\sin x)^{2y-1}dx$, show that, if $0 < c \leqq y$,

$$\frac{df(y)}{dy} = 2\int_0^{\frac{\pi}{2}} (\sin x)^{2y-1}\log(\sin x)dx.$$

The integral $f(y)$ converges uniformly if $y \geqq c > 0$, and so does the second integral (§ 146, Exs. 4, 16), so that differentiation is legitimate; similarly the 2nd, 3rd ... derivatives of $f(y)$ are found by differentiating under the integral sign.

Now $\qquad f(y) = \frac{1}{2}\sqrt{\pi}\Gamma(y)/\Gamma(y+\frac{1}{2});$

therefore $\qquad \dfrac{df(y)}{dy} = \frac{1}{2}\sqrt{\pi}\dfrac{\Gamma(y)}{\Gamma(y+\frac{1}{2})}\{\psi(y-1) - \psi(y-\frac{1}{2})\},$

where $\psi(y)$ is the derivative of $\log\Gamma(y+1)$ (§ 97).

Ex. 2. If $f(y) = \int_0^y \frac{\psi(x)dx}{\sqrt{(y-x)}}$, $\quad < c \leqq y$, where $\psi(x)$ is a continuous function of x, find $df(y)/dy$.

Here, as happens not infrequently, the integral obtained by differentiating under the integral sign does not converge; the integral for $f(y)$ must therefore be transformed in some way (compare *E.T.* p. 468, Ex. 4). In this case change the variable of integration from x to t where $x = yt$; then

$$f(y) = \int_0^1 y^{\frac{1}{2}}\psi(yt)\frac{dt}{(1-t)^{\frac{1}{2}}},$$

and it is now legitimate to differentiate under the integral sign.

Thus
$$\frac{df(y)}{dy}=\int_0^1\left\{\frac{\psi(yt)}{2y^{\frac{1}{2}}}+y^{\frac{1}{2}}\,t\,\psi'(yt)\right\}\frac{dt}{(1-t)^{\frac{1}{2}}},$$
or, in terms of x as variable of integration,
$$\frac{df(y)}{dy}=\int_0^y\frac{\psi(x)+2x\psi'(x)}{2y(y-x)^{\frac{1}{2}}}dx.$$
It should be noted that, as in this example, a change of the variable of integration is often useful for reducing an integral with variable limits to one with constant limits.

156. Evaluation of Improper Double Integrals. The General Theorems stated for the proper double integral in § 125 are true for the improper integral, as may be seen by applying the definition. Theorem III may be an exception, when $F_1(x, y)$ and $F_2(x, y)$ have common singularities (cf. § 145, Theorem A). It should be remembered that the improper double integral of $F(x, y)$ is always understood to be *absolutely* convergent (§ 150).

The evaluation of an improper double integral is usually effected in practice by means of a repeated integral, and the general method of procedure may be illustrated by the Problem of § 149; see that section for the notation.

When the auxiliary curves have been drawn so as to exclude the points and curves of discontinuity of $F(x, y)$ from the area T, the proper double integral of $F(x, y)$ over the contracted area T' is expressed by equation (1) as a repeated integral
$$\iint_{T'}F(x, y)\,dx\,dy=\int_\delta^{c-\delta'}dx\int_{\delta'}^{c-x}F(x, y)dy, \quad \text{.................(1)}$$
where $F(x, y)=x^{m-1}y^{n-1}f(x, y)$, $0<m<1$, $0<n<1$, $f(x, y)$ being continuous in T.

The next step is to show that when δ and δ' tend to zero the repeated integral in (1) tends to the integral
$$\int_0^c dx\int_0^{c-x}F(x, y)dy, \quad \text{............................(2)}$$
where the integral in (2) is taken for the given area T.

Now the integral $g(x)$, where
$$g(x)=\int_0^{c-x}x^{m-1}y^{n-1}f(x, y)dy$$
converges uniformly for the range $0<\delta\leqq x\leqq c$, and therefore the *limit* for δ' tending to zero of the repeated integral in (1) will not be altered by taking zero instead of δ' as the lower limit of the integral with respect to y. Next, when δ and δ' tend to zero, the integral
$$\int_\delta^{c-\delta'}dx\int_0^{c-x}x^{m-1}y^{n-1}f(x, y)dy=\int_\delta^{c-\delta'}g(x)dx$$
tends to the integral (2) because the integral of $g(x)$ is convergent. It

follows, therefore, that when δ and δ' tend to zero the double integral also tends to a limit and that limit is equal to the repeated integral (2).

The discussion of repeated integrals in the preceding articles will probably be sufficient to enable the student to evaluate an improper double integral in such cases as usually occur. For fuller information on the general problem he may consult the memoir of De la Vallée Poussin in *Journ. de Math.* (see § 151) or Stolz's *Differential- und Integralrechnung*, vol. 3.

One general theorem may be stated.

THEOREM. *If $F(x, y)$ does not change sign in an area A, or if A can be divided into a finite number of areas in each of which $F(x, y)$ does not change sign, the double integral of $F(x, y)$ over A will converge and will be equal to the repeated integral U (or V) provided the repeated integral converges.*

157. Multiple Integrals. Change of Variables. The extension to integrals of functions of more than two variables of the definition of the integral when the integrand is not bounded in the field of integration or the field not finite is made in the same way as for a double integral and hardly requires further explanation. The evaluation of the improper multiple integral is usually effected by means of repeated integrals, and the principles elucidated in the consideration of repeated integrals for two variables are to be applied when there are more than two variables. When the change of order of integration is made in the manner shown in the examples of § 133, the principles that underlie the change in the case of two variables come constantly into play.

Again, when a change in the variables of integration is to be made, the conditions regarding the one-to-one correspondence between the old and new variables and the continuity of the variables (§ 134) must be satisfied. If, when the variables x, y are changed to u, v, the area A is changed to the area B, let A' and B' be the corresponding contracted areas for which a proper double integral of $F(x, y)$ and of the transformed function $F_1(u, v)$ exists; then

$$\iint_{A'} F(x, y)\, dx\, dy = \iint_{B'} F_1(u, v)\, |\, J\, |\, du\, dv,$$

and if the double integral of $F(x, y)$ over A' tends to a limit

when A' tends to A, the double integral of $F_1(u, v)\,|J\,|$ will also tend to a limit when B' tends to B, and the two limits will be equal.

The following examples illustrate various cases.

Ex. 1. Transform the integral of the function $F(x, y)$ of § 149 by the change of variables $x+y=u$, $x=uv$.

The contracted area T', which is bounded by the lines $x=\delta$, $y=\delta'$, and $x+y=c$, becomes the area B' in the plane of the co-ordinates u, v, bounded by $u=c$, and tho hyperbolas $v=\delta/u$, $v=1-(\delta'/u)$, which intersect at the point $u=\delta+\delta'$, $v=\delta/(\delta+\delta')$. Hence, if $F_1(u, v)$ is the value of $F(x, y)$ in terms of u and v,

$$\iint_{T'} F(x, y)\,dx\,dy = \iint_{B'} F_1(u, v)\,u\,du\,dv. \quad \text{...............(i)}$$

Now $F(x, y)=x^{m-1}y^{n-1}f(x, y)$ and therefore

$$\iint_{B'} F_1(u, v)\,u\,du\,dv = \int_{\delta+\delta'}^{c} u^{m+n-1}du \int_{\frac{\delta}{u}}^{1-\frac{\delta'}{u}} v^{m-1}(1-v)^{n-1}f_1(u, v)dv, \quad \text{..(ii)}$$

where $f_1(u, v)$ is the value of $f(x, y)$ in terms of u and v.

Let
$$g(u)=u^{m+n-1}\int_0^1 v^{m-1}(1-v)^{n-1}f_1(u, v)dv\,;$$
then $g(u)$ converges uniformly for the range $\delta+\delta' \leqq u \leqq c$, since $0<m$, $0<n$, so that the *limit* for δ and δ' tending to zero of the repeated integral in (ii) will not be changed if in the limits of the integral with respect to v the numbers δ and δ' are made zero. Next, the integral of $g(u)$ converges when δ and δ' tend to zero, and therefore

$$\iint_T F(x, y)\,dx\,dy = \iint_B F_1(u, v)\,u\,du\,dv$$
$$= \int_0^c u^{m+n-1}du \int_0^1 v^{m-1}(1-v)^{n-1}f_1(u, v)dv. \quad \text{.......(iii)}$$

The repeated integral in (iii) is the same as that obtained in § 133, Ex. 1.

Ex. 2. Prove, by the same transformation as in Ex. 1, that

$$\int_0^1 dx\int_0^{1-x} x^{m-1}y^{n-1}(1-x-y)^{p-1}dy$$
$$=\int_0^1 u^{m+n-1}(1-u)^{p-1}du\int_0^1 v^{m-1}(1-v)^{n-1}dv,$$

where m, n, p are all positive.

See § 133, Ex. 2. The other examples in § 133 can all be dealt with in the same way.

Ex. 3. Prove tnat $\Gamma(m)\Gamma(n)=\Gamma(m+n)B(m, n)$ if $m>0$, $n>0$.

If a and b are any two positive numbers and R the rectangle $(0, 0\,;\,a, b)$, then, $0<m$, $0<n$,

$$\iint_R e^{-(x+y)}x^{m-1}y^{n-1}dx\,dy = \int_0^a e^{-x}x^{m-1}dx\int_0^b e^{-y}y^{n-1}dy. \quad \text{.........(i)}$$

When a and b tend (independently) to ∞ the repeated integral in (i) tends to the product $\Gamma(m)\Gamma(n)$.

Next, whatever be the relative magnitudes of a and b, it is possible to find two triangles T and T', bounded respectively by the lines

$$x=0,\ y=0,\ x+y=c, \text{ and } x=0,\ y=0,\ x+y=c'>c,$$

such that R lies within T' and T within R. Hence, since the integrand of the double integral is positive,

$$\iint_T e^{-(x+y)}x^{m-1}y^{n-1}dx\,dy<\iint_R e^{-(x+y)}x^{m-1}y^{n-1}dx\,dy$$
$$<\iint_{T'} e^{-(x+y)}x^{m-1}y^{n-1}dx\,dy.$$

If these integrals are transformed by the substitution $x+y=u,\ x=uv$, they give the inequalities

$$B(m,n)\int_0^c e^{-u}u^{m+n-1}du<\iint_R e^{-(x+y)}x^{m-1}y^{n-1}dx\,dy<B(m,n)\int_0^{c'} e^{-u}u^{m+n-1}du,$$

and therefore when c and c' tend to ∞ the double integral in (i) tends to $B(m,n)\,\Gamma(m+n)$. The required relation is thus established.

Ex. 4. If T is the triangle bounded by the lines $x=0$, $y=0$, and $x+y=1$, and if I is the integral, $0<m, 0<n$,

$$I=\iint_T x^{m-1}y^{n-1}dx\,dy,$$

prove that $\quad I=B(m,n)/(m+n)=\Gamma(m)\Gamma(n)/\Gamma(m+n+1).$

We have $\quad I=\int_0^1 x^{m-1}dx\int_0^{1-x}y^{n-1}dy=\frac{1}{n}\int_0^1 x^{m-1}(1-x)^n\,dx.$

Also $\quad I=\int_0^1 y^{n-1}dy\int_0^{1-y}x^{m-1}dx=\frac{1}{m}\int_0^1 y^{n-1}(1-y)^m dy,$

or $\quad I=\frac{1}{m}\int_0^1 x^m(1-x)^{n-1}dx,$

by the substitution $y=1-x$. Hence

$$(m+n)\,I=\int_0^1 x^{m-1}(1-x)^{n-1}dx=B(m,n).$$

Ex. 5. If $F(x,y)=f(x,y)/(x^2+y^2)^n$, where $f(x,y)$ is continuous, show that the double integral of $F(x,y)$ over any (finite) area A within which the origin lies is convergent provided $n<1$.

Change to polar co-ordinates; then

$$\iint_A F(x,y)\,dx\,dy=\iint_A \frac{f(x,y)}{r^{2n}}r\,dr\,d\theta=\iint_A f(x,y)r^{1-2n}dr\,d\theta.$$

Take a circle of radius ε, with centre at the origin; in this circle $|f(x,y)|<K$, a constant, and therefore the integral over this area is numerically less than $\pi K\varepsilon^{2-2n}/(1-n)$, an expression which tends to zero when ε tends to zero if $n<1$.

Ex. 6. If $F(x,y,z)=f(x,y,z)/r^n$, where $f(x,y,z)$ is a continuous function of the variables x, y, z and $r^2=(x-a)^2+(y-b)^2+(z-c)^2$, show

that the triple integral of $F(x, y, z)$ taken throughout any (finite) field within which the point (a, b, c) lies is convergent if $n < 3$.

Take polar co-ordinates with the point (a, b, c) as origin; then the integral, taken throughout a sphere of radius ε with the point (a, b, c) as centre, is numerically less than $(|f(x, y, z)| < K)$

$$K\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta \int_0^{\varepsilon} r^{2-n} dr = \frac{4\pi K \varepsilon^{3-n}}{3-n},$$

and therefore tends to zero when $\varepsilon \to 0$ if $n < 3$.

Ex. 7. If A is the area in the first quadrant, bounded by the x-axis, the curve $y = x^\lambda$, $\lambda > 1$, and the circle $r = 1$, prove that the double integral of $1/(x^2 + y^2)^n$ over the area A is convergent provided $n < (1 + \lambda)/2$.

(Stolz, *l.c.* p. 194.)

Ex. 8. If $F(x, y) = f(x, y)/(a^2 - x^2 - y^2)^n$, where $f(x, y)$ is a continuous function of x and y, show that the double integral of $F(x, y)$ over the circle $x^2 + y^2 = a^2$ is convergent if $n < 1$.

If the area of integration is the triangle bounded by the lines $x = 0$, $y = 0$, $x + y = a$, show that the double integral converges if $n < 2$.

(Stolz, *l.c.* p. 197.)

In the second case draw the auxiliary line $x = a - \delta$, where δ is positive and arbitrarily small, so as to cut off the corner of the triangle at $(a, 0)$; it is not hard to show that the integral over this small area tends to zero with δ if $n < 2$. By symmetry the corresponding integral at the corner $(0, a)$ also tends to zero if $n < 2$, and therefore the double integral converges.

EXERCISES XVIII.

1. $\displaystyle\int_0^1 \log\left(\frac{1+ax}{1-ax}\right)\frac{dx}{x\sqrt{(1-x^2)}} = \pi \sin^{-1} a,\ a^2 \leqq 1.$

2. $\displaystyle\int_0^\theta \log(1 + \tan\theta \tan x)dx = \theta \log\sec\theta,\quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}.$

3. $\displaystyle\int_0^1 \log\left\{\frac{(1+ax)^{\frac{1}{x}}}{1+a}\right\}\frac{dx}{1-x} = \tfrac{1}{2}\{\log(1+a)\}^2,\ a > 0.$

4. If $\displaystyle u = \int_0^\pi \frac{\sin^{2n} x\, dx}{(1 - 2a\cos x + a^2)^n}$, where $-1 < a < 1$ and n is a positive integer, prove that

(i) $\displaystyle\frac{d}{da}\left(a^{2n+1}\frac{du}{da}\right) = 0;$

(ii) $\displaystyle u = \frac{1 . 3 . 5 . \ldots (2n-1)}{2 . 4 . 6 \ldots 2n}\pi.$

(Poisson.)

What is the value of the integral if $|a| > 1$?

5. If $\displaystyle y = x\int_0^{\frac{\pi}{2}} \cos(x^2 \sin\theta)\sqrt{(\cos\theta)}\,d\theta$, prove that $\displaystyle\frac{d^2y}{dx^2} + 4x^2 y = 0.$

6. Apply Ex. 1 of § 155 to prove the following results:

(i) $\int_0^{\frac{\pi}{2}} \sin x \log(\sin x)\, dx = \log 2 - 1$;

(ii) $\int_0^{\frac{\pi}{2}} \sin x\, (\log \sin x)^2\, dx = (\log 2 - 1)^2 + 1 - \frac{\pi^2}{12}$;

(iii) $\int_0^{\frac{\pi}{2}} (\log \sin x)^2\, dx = \frac{\pi}{2}\left\{(\log 2)^2 + \frac{\pi^2}{12}\right\}$;

(iv) $\int_0^{\frac{\pi}{2}} \frac{\log(\sin x)}{\sqrt{(\sin x)}}\, dx = -\frac{\sqrt{\pi}}{4\sqrt{2}}\{\Gamma(\tfrac{1}{4})\}^2$;

(v) $\int_0^{\frac{\pi}{2}} \sqrt{(\sin x)} \log(\sin x)\, dx = \pi(\pi - 4)\sqrt{(2\pi)}/\{\Gamma(\tfrac{1}{4})\}^2$.

7. If $u = \int_0^{\frac{\pi}{2}} (\sin t)^{2x-1}(\cos t)^{2y-1}\, dt$, where $x > 0$ and $y > 0$, show that the derivatives of u with respect to x and y may be obtained by differentiating under the integral sign.

Deduce that

(i) $\int_0^{\frac{\pi}{2}} \log(\sin t) \log(\cos t)\, dt = \frac{\pi}{2}\left\{(\log 2)^2 - \frac{\pi^2}{24}\right\}$;

(ii) $\int_0^{\frac{\pi}{2}} \sin t \log(\sin t) \log(\cos t)\, dt = 2 - \log 2 - \frac{\pi^2}{8}$.

8. Prove that $\int_0^1 dx \int_x^1 \frac{y\, dy}{(1 + xy)^2(1 + y^2)} = \frac{\pi - 1}{4}$.

9. Prove (i) $\int_0^a dx \int_0^x \frac{y^2\, dy}{\sqrt{\{(a - x)(x - y)\}}} = \frac{\pi}{3} a^3$;

(ii) $\int_0^a dx \int_0^x \frac{\sec^2 y\, dy}{\sqrt{\{(a - x)(x - y)\}}} = \pi \tan a$.

10. Show that, if $0 < n < 1$,

$$\int_a^b \frac{dy}{(b - y)^{1-n}} \int_a^y \frac{f'(x)\, dx}{(y - x)^n} = \frac{\pi}{\sin n\pi}\{f(b) - f(a)\}.$$

11. Prove that

$$\int_0^a \frac{dx}{(a - x)^{\frac{2}{3}}} \int_0^x \frac{dy}{(x - y)^{\frac{2}{3}}} \int_0^y \frac{f'(z)\, dz}{(y - z)^{\frac{2}{3}}} = [\Gamma(\tfrac{1}{3})]^3\{f(a) - f(0)\}.$$

12. Prove that, if n is a positive integer,

$$\int_0^a \frac{dx_1}{(a - x_1)^{\frac{n-1}{n}}} \int_0^{x_1} \frac{dx_2}{(x_1 - x_2)^{\frac{n-1}{n}}} \cdots \int_0^{x_{n-1}} \frac{f'(x_n)\, dx_n}{(x_{n-1} - x_n)^{\frac{n-1}{n}}}$$
$$= \left\{\Gamma\left(\frac{1}{n}\right)\right\}^n \left\{f(a) - f(0)\right\}.$$ (Tait.)

13. If $F'(x)$ denote $dF(x)/dx$, show that

$$\int_0^1 \frac{d\lambda}{\sqrt{(1-\lambda^2)}} \int_0^1 \frac{F'(k\lambda\mu)\,\mu\,d\mu}{\sqrt{(1-\mu^2)}} = \frac{\pi}{2k}\{F(k)-F(0)\}.$$

14. $\displaystyle\int_0^1 y^{\frac{3}{2}}(1-y)^{\frac{1}{2}}\,dy \int_0^1 \frac{x^2(1-x)\,dx}{(1+x+xy)^5} = \frac{\pi}{576\sqrt{6}}.$

15. If $P = \displaystyle\int_0^1\int_0^1 \frac{dx\,dy}{1-xy}$, $Q = \displaystyle\int_0^1\int_0^1 \frac{dx\,dy}{1+xy}$, prove that both integrals are convergent and then show that

(i) $P - Q = \frac{1}{2}P$, or $P = 2Q$;

(ii) $P + Q = \displaystyle\int_0^1 dx \int_{-1}^1 \frac{dy}{1+xy}$;

(iii) $\displaystyle\int_{-1}^1 \frac{dy}{1+xy} = \int_{-1}^1 \frac{du}{1+2xu+x^2}$, $(1+xy)^2 = 1+2xu+x^2$;

and deduce from (ii) and (iii) that $P+Q = \dfrac{\pi^2}{4}$, so that

$$P = \frac{\pi^2}{6},\quad Q = \frac{\pi^2}{12}.$$

[*Archiv. der Math. und Phys.*, vol. 13, p. 362, Year 1908.]

16. One loop of the curve $r^2\cos^2\theta = a^2\cos 2\theta$ makes a complete revolution about the initial line; prove that the volume enclosed by the surface generated is $\frac{1}{6}\pi(10-3\pi)a^3$.

17. The sphere $x^2+y^2+z^2=1$ is pierced by the cylinder

$$2x^2(x^2+y^2) = 3(x^2-y^2);$$

show that the area of the spherical surface that lies inside the cylinder is

$$2\pi + 8\sqrt{2}\log(1+\sqrt{2}) - 4\sqrt{6}\log(\sqrt{3}+\sqrt{2}).$$

18. Show that the area enclosed by the curve

$$\left(\frac{x}{a}+\frac{y}{b}\right)^{2n} + \left(\frac{x}{b}-\frac{y}{a}\right)^{2n} = 1$$

is

$$\frac{1}{n}\cdot\frac{\left[\Gamma\left(\frac{1}{2n}\right)\right]^2}{\Gamma\left(\frac{1}{n}\right)}\cdot\frac{a^2b^2}{a^2+b^2}.$$

Note. Further examples may be found in Exercises XVI, where, as in Exs. 13, 14, 15, the indices $l, m, \ldots$ may have values for which the integrals converge though the integrand is not bounded.

CHAPTER XIV

DOUBLE INTEGRALS: RANGE INFINITE

158. Range of Integration Infinite. The discussion of the integrals when the range of integration is infinite is based on the methods of De la Vallée Poussin (see § 151), and is to a large extent identical with that given in Chapter XXI of the *Elementary Treatise*; it is, however, more general, as the integrand $F(x, y)$ may have infinite discontinuities such as are specified in §§ 151, 152.

Definition. The integral

$$\int_a^\infty F(x, y)dx \text{ ....(A)}$$

is said to converge uniformly for the range $a' \quad y \leqq b'$ if (i) the integral

$$\int_a^B F(x, y)dx$$

converges uniformly for every fixed value of B, and if (ii), when

$$R_b = \int_b^\infty F(x, y)dx,$$

B can be chosen so that $|R_b| < \varepsilon$ if $b \geqq B$, whatever value in the range $a' \leqq y \leqq b'$ is assigned to y, ε being any given arbitrarily small positive number.

If conditions (i) and (ii) are satisfied for every value of y greater than or equal to a', the integral (A) is said to converge uniformly for the *unlimited* range $y' \geqq a'$.

If conditions (i) and (ii) are satisfied for the range $a' \leqq y \leqq b'$, where b' is any *fixed* number, no matter how large it may be, the integral (A) is said to converge uniformly in an *arbitrarily large* interval (a', b') or for an *arbitrarily large* range $a' \leqq y \leqq b'$. (See *E.T.* p. 462, top of page.)

The integral (A) is said to *converge uniformly in general* in an

interval (a', b') if there is a finite number of values of y in the interval for which the convergence ceases to be uniform. If there is only one such number c', where $a'<c'<b'$, and if δ and δ' are arbitrarily small positive numbers, the convergence is uniform in the intervals $(a', c'-\delta)$ and $(c'+\delta', b')$. Compare § 151.

The condition that the integral (A) should converge for a given value y_1 of y is

$$\left|\int_b^c F(x, y_1)dx\right| < \varepsilon \text{ if } c>b \geqq B;$$

the integral converges uniformly for a given range of y if B is such that the inequality is satisfied for *every* value y_1 of y in the range. The tests for convergence, stated and illustrated in Chapter XII, are now to be used as tests for uniform convergence, the new element being that B must be independent of y. Three principal forms will now be stated, the proofs being left to the student, as, after what has been already done, they offer no difficulty; the interval for y may be (a', b') or (a', ∞).

The M-Test. The conditions are: (i) $|F(x, y)| \leqq M(x)$, if $x \geqq a$, where $M(x)$ is positive and independent of y, and (ii) the integral

$$\int_a^\infty M(x)dx$$

is convergent.

Cor. If $F(x, y) = \varphi(x)\psi(x, y)$, where $|\psi(x, y)| < K$, a constant, for every value of y in the range, and if

$$\int_a^\infty |\varphi(x)|dx$$

converges, then the integral (A) converges uniformly.

Abel's Test. $F(x, y)$ is a product, $\varphi(x, y)\psi(x)$, where $\psi(x)$ is independent of y. The conditions are: (i) $\varphi(x, y)$, when y is constant, is a positive, monotonic, decreasing function of x; (ii) $\varphi(a, y)$ is less than a constant K for every value of y in the range; (iii) the integral

$$\int_a^\infty \psi(x)dx$$

is convergent.

For the proof compare § 145, Theorem E.

Cor. If $F(x, y)=\varphi(x, y)\,\psi(x, y)$, so that ψ is a function of y as well as of x, the test will hold, provided that, in place of (iii), is substituted the condition that the integral

$$\int_a^\infty \psi(x, y)\,dx$$

converges *uniformly*.

Dirichlet's Test. Here also $F(x, y)=\varphi(x, y)\,\psi(x)$. The conditions are: (i) $\varphi(x, y)$, when y is constant, is a positive, monotonic, decreasing function of x; (ii) $\varphi(x, y)$ tends to zero when x tends to infinity for every value of y in the range; (iii) the integral

$$\int_a^\infty \psi(x)\,dx$$

oscillates finitely.

In applying any test, advantage should be taken of any suitable transformation of the integral; for example, by using integration by parts or by expressing the integral as the sum of two or more integrals.

The following examples are drawn chiefly from the Memoir of De la Vallée Poussin in the *Ann. ... de Bruxelles*.

Ex. 1. $$\int_0^\infty \frac{\sin(xy)}{x}\,dx \text{ and } \int_0^\infty \frac{\sin y \sin(xy)}{x}\,dx.$$

The first integral converges uniformly for $y \geqq a' > 0$ or for $y \leqq a' < 0$; the second converges uniformly for every value of y, since

$$\left|\int_b^c \frac{\sin y \sin(xy)}{x}\,dx\right| < 2\left|\frac{\sin y}{y}\right| \leqq \frac{2}{b}.$$

Ex. 2. If $F(x, y)=e^{-cy}\cos x \sin(xy)/x$, $c>0$, and if

$$f(y)=\int_0^\infty F(x, y)\,dx \text{ and } g(x)=\int_0^\infty F(x, y)\,dy,$$

prove that $f(y)$ converges uniformly in general for the range $y \geqq 0$ and $g(x)$ converges uniformly for the range $x \geqq 0$.

$$e^{-cy}\int_b^\infty \frac{\cos x \sin(xy)}{x}\,dx=\frac{e^{-cy}}{b}\int_b^\xi \cos x \sin(xy)\,dx,\ \xi>b,$$
$$=\frac{e^{-cy}}{2b}\left\{\frac{\cos(y+1)b-\cos(y+1)\xi}{y+1}+\frac{\cos(y-1)b-\cos(y-1)\xi}{y-1}\right\}.$$

The convergence therefore ceases to be uniform for $y=1$ and for that value only. The convergence of $g(x)$ is still more easily tested.

Ex. 3. If $F(x, y)=e^{-cy}\sin x \cos(xy)/x$, $c>0$, and if $f(y)$ and $g(x)$ denote the same integrals as in Ex. 2, with the new value of $F(x, y)$, prove that $f(y)$ and $g(x)$ converge as in Ex. 2.

Ex. 4. If $v(\eta)=\int_0^\infty \frac{\cos x}{x}\,dx \int_0^\eta e^{-cy}\sin(xy)\,dy$, $c>0$, prove that $v(\eta)$ converges uniformly for the unlimited range $\eta \geqq 0$.

Here $v(\eta)$ is equal to

$$\int_0^\infty \frac{\cos x(1-e^{-c\eta}\cos \eta x)}{x^2+c^2}\,dx - ce^{-c\eta}\int_0^\infty \frac{\sin \eta x}{x}\cdot\frac{\cos x}{x^2+c^2}\,dx,$$

and each of these integrals converges uniformly for the range $\eta \geqq 0$.

Ex. 5. If $v(\eta)=\int_0^\infty \frac{\sin x}{x}\,dx\int_0^\eta e^{-cy}\cos(xy)\,dy$, $c>0$, show, as in Ex. 4, that $v(\eta)$ converges uniformly for the range $\eta \geqq 0$.

Ex. 6. If $v(\eta)=\int_0^\infty e^{ix}dx\int_0^\eta e^{-xy^2}dy$, show that $v(\eta)$ converges uniformly for the range $\eta \geqq 0$.

Let
$$\varphi(x, \eta)=\int_0^\eta e^{-xy^2}dy;$$
then, η being constant, $\varphi(x, \eta)$ is a positive, monotonic, decreasing function of x and tends to zero when $x \to \infty$ for every value of η such that $\eta \geqq 0$. Now apply Dirichlet's Test.

Ex. 7. If $f(y)=\int_0^\infty \frac{dx}{x}\int_1^y \Gamma(t)\{e^{-x}-(1+x)^{-t}\}dt$, show that $f(y)$ converges uniformly for the range, $1 \leqq y \leqq b'$, where b' is arbitrarily large.

Let $\varphi(x, t)=e^{-x}-(1+x)^{-t}$, $t \geqq 1$; then $\varphi(x, t)/x \to (t-1)$ when $x \to 0$, so that the integral converges at the lower limit $x=0$.

Again, if $1 \leqq t \leqq b'$, where b' is arbitrarily large, the function $\Gamma(t)\varphi(x, t)/x$ is a continuous, bounded function of x and t and therefore, by the First Theorem of Mean Value,

$$\int_1^{b'}\frac{\Gamma(t)\varphi(x, t)}{x}\,dt=(b'-1)\frac{\Gamma(t')\varphi(x, t')}{x},\quad 1 \leqq t' \leqq b'.$$

Hence
$$f(y)=(b'-1)\int_0^\infty \Gamma(t')\left\{e^{-x}-(1+x)^{-t'}\right\}\frac{dx}{x}.$$

Now t' depends on x; but $\Gamma(t') \leqq \Gamma(b')$ and

$$\left|R_b\right|=(b'-1)\left|\int_b^\infty \Gamma(t')\left\{e^{-x}-(1+x)^{-t'}\right\}\frac{dx}{x}\right|$$
$$<(b'-1)\Gamma(b')\left\{\int_b^\infty \frac{e^{-x}}{x}\,dx+\int_b^\infty \frac{dx}{x(1+x)}\right\},$$

so that, by choice of b, $|R_b|$ can be made arbitrarily small whatever value y may have in the range $1 \leqq y \leqq b'$. (See also § 160, Ex. 1.)

Ex. 8. $f(y)=\int_0^\infty x\sin(x^3-xy)\,dx.$

The coefficient x of $\sin(x^3-xy)$ may be expressed in the form

$$x=\frac{3x^2-y}{3x}+\frac{y}{9}\frac{3x^2-y}{x^3}+\frac{y^2}{9}\frac{1}{x^3},$$

and therefore

$$\int_b^\infty x\sin(x^3-xy)dx$$
$$=\int_b^\infty \frac{(3x^2-y)\sin(x^3-xy)}{3x}\,dx+\frac{y}{9}\int_b^\infty \frac{(3x^2-y)\sin(x^3-xy)}{x^3}\,dx$$
$$+\frac{y^2}{9}\int_b^\infty \frac{\sin(x^3-xy)}{x^3}\,dx.$$

The integral of $(3x^2-y)\sin(x^3-xy)$ is $-\cos(x^3-xy)$, and therefore, by applying the Second Theorem of Mean Value, we see that the first two integrals tend to zero when $b\to\infty$, and obviously the third integral also tends to zero when $b\to\infty$. Hence $f(y)$ converges uniformly in an arbitrarily large interval $(0, b')$.

Ex. 9. The integral $\int_0^\infty e^{-x}x^{n-1}(\log x)^m dx$ converges uniformly, if m is a positive integer and $n \geqq c > 0$.

The result follows at once from Ex. 17 of § 146.

159. Continuity of Integrals. The discontinuities of the integrand $F(x, y)$, when the ranges of x and y are finite, are understood to satisfy the conditions stated in § 152.

THEOREM I. *If the integral $f(y)$, where*

$$f(y) = \int_a^\infty F(x, y)\,dx,$$

converges uniformly for the range $a' \leqq y \leqq b'$, it is a continuous function of y for that range.

Let $\qquad R_b = \int_b^\infty F(x, y)\,dx$; then

$$f(y) = \int_a^b F(x, y)\,dx + R_b = f_1(y) + R_b, \text{ say.}$$

It is possible, since $f(y)$ converges uniformly, to choose b so that, when $a' \leqq y \leqq b'$, we shall have $|R_b| < \varepsilon$, where ε has the usual meaning; b, when so chosen, is to be kept fixed.

Next, by § 152, Theorem I, the integral $f_1(y)$ is a continuous function of y for the range $a' \leqq y \leqq b'$, and therefore, if y_1 and y_2 are any two values in that range, there is a number η such that $|f_1(y_1) - f_1(y_2)| < \varepsilon$ if $|y_1 - y_2| < \eta$. Hence, if $|y_1 - y_2| < \eta$,

$$|f(y_1) - f(y_2)| \leqq |f_1(y_1) - f_1(y_2)| + 2|R_b| < 3\varepsilon,$$

so that $f(y)$ is a continuous function of y for the range $a' \leqq y \leqq b'$.

The following theorem is an extension of Theorem II of § 152 and is required later; $f(y)$ and $g(x)$ denote the integrals

$$f(y) = \int_a^\infty F(x, y)\,dx, \quad g(x) = \int_{a'}^{b'} F(x, y)\,dy.$$

THEOREM II (i). *If the integral $f(y)$ converges only uniformly in general for the range $a' \leqq y \leqq b'$, while the integral $g(x)$ converges*

uniformly for the arbitrarily large range $a \leqq x \leqq b$, *then* $v(\eta)$, *where*

$$v(\eta)=\int_a^{\infty} dx \int_{a'}^{\eta} F(x, y)dy,$$

is a continuous function of η *for the range* $a' \leqq \eta \leqq b'$, *provided the integral converges uniformly for that range.*

Let
$$R_b=\int_b^{\infty} dx \int_{a'}^{\eta} F(x, y)dy;$$
then, since $v(\eta)$ converges uniformly, b may be chosen so that, if $a' \leqq \eta \leqq b'$, we shall have $|R_b|<\varepsilon$. When thus chosen, b is to be kept fixed; thus

$$\begin{aligned} v(\eta) &= \int_a^{b} dx \int_{a'}^{\eta} F(x, y)dy + R_b, \quad |R_b|<\varepsilon, \\ &= w(\eta)+R_b, \text{ say.} \end{aligned}$$

Now, by § 153, Theorem II,

$$w(\eta)=\int_{a'}^{\eta} dy \int_a^{b} F(x, y)dx,$$

and therefore $w(\eta)$ is a continuous function of η for the range $a' \leqq \eta \leqq b'$. Hence it follows, by the same reasoning as in the proof of Theorem I, that $v(\eta)$ is continuous for the range $a' \leqq \eta \leqq b'$.

THEOREM II (ii). *If the integral* $f(y)$ *converges uniformly for the range* $a' \leqq y \leqq b'$, *while the integral* $g(x)$ *only converges uniformly in general for the arbitrarily large range* $a \leqq x \leqq b$, *then the integral* $v(\eta)$ *is a continuous function of* η *for the range* $a' \leqq \eta \leqq b'$, *provided* $v(\eta)$ *converges uniformly for that range.*

The proof is the same as for Theorem II (i).

Cor. If $F(x, y)$ does not change sign it may be shown, as in the proof of § 152, Theorem III, that the continuity of $v(\eta)$ follows from the existence of $v(b')$.

THEOREM III. *Let* $\psi(\eta)=\int_a^{\infty} f(x, \eta)dx$. *Given* (i) *that the integral* $\psi(\eta)$ *converges uniformly for the unlimited range* $\eta \geqq a'$, *and* (ii) *that, when* $\eta \to \infty$, *the function* $f(x, \eta)$ *converges uniformly to* $\varphi(x)$ *for the arbitrarily large range* $a \leqq x \leqq b$; *then*

$$\mathop{\mathcal{L}}_{\eta\to\infty} \int_a^{\infty} f(x, \eta)dx = \int_a^{\infty} \varphi(x)dx = \int_a^{\infty} \left[\mathop{\mathcal{L}}_{\eta\to\infty} f(x, \eta) \right] dx.$$

The following proof is that given in the *Elementary Treatise* (pp. 465-6).

(I). $\psi(\eta)$ tends to a limit. Let η' and η'' be any two values of η; then the difference $\{\psi(\eta')-\psi(\eta'')\}$ is equal to

$$\int_a^b\{f(x,\eta')-f(x,\eta'')\}\,dx+\int_b^\infty f(x,\eta')\,dx-\int_b^\infty f(x,\eta'')\,dx\ \ldots\ldots(1)$$

$=\lambda+\mu-\nu$, say.

By condition (i), b may be chosen so that $|\mu|<\varepsilon$, $|\nu|<\varepsilon$, whatever values η' and η'' may take in the interval (a', ∞); when b has been so chosen it is to be kept fixed.

By condition (ii) there is a number Y such that $|f(x,\eta')-f(x,\eta'')|$ is less than $\varepsilon/(b-a)$ if $\eta'>Y$, and $\eta''>Y$, and therefore $|\lambda|<\varepsilon$ if $\eta'>Y$ and $\eta''>Y$.

Hence $|\psi(\eta')-\psi(\eta'')|<3\varepsilon$ if $\eta'>Y$ and $\eta''>Y$, so that $\psi(\eta)$ tends to a limit when $\eta\to\infty$. Let the limit be denoted by P; it has next to be proved that

$$P=\int_a^\infty\varphi(x)\,dx.$$

(II). We have, b being for the moment undetermined,

$$\int_a^b\varphi(x)\,dx-P=\int_a^b\{\varphi(x)-f(x,\eta)\}\,dx-\int_b^\infty f(x,\eta)\,dx$$
$$+\left[\int_a^\infty f(x,\eta)\,dx-P\right]\ldots\ldots\ldots(2)$$

$=\alpha-\beta+\gamma$, say.

P is the limit of $\psi(\eta)$, and therefore Y_1 may be chosen so that $|\gamma|<\varepsilon$ if $\eta>Y_1$. By condition (i), B may be chosen so that $|\beta|<\varepsilon$ if $b\geqq B$, whatever be the value of η; choose such a value of b and then keep it fixed. By condition (ii), Y_2 may be chosen so that $|\alpha|<\varepsilon$ if $\eta>Y_2$. Hence, if $\eta>Y_3$, where Y_3 is greater than either Y_1 or Y_2, and $b\geqq B$, we have

$$\left|\int_a^b\varphi(x)\,dx-P\right|<3\varepsilon;$$

but the expression on the left of this inequality is independent of η, and therefore

$$\left|\int_a^b\varphi(x)\,dx-P\right|<3\varepsilon,\text{ if } b\geqq B;$$

that is,
$$\mathop{\mathcal{L}}_{b\to\infty}\int_a^b\varphi(x)\,dx=P,\quad\text{or}\quad\int_a^\infty\varphi(x)\,dx=P.$$

The theorem is therefore proved.

The special use of the theorem occurs when $f(x, \eta)$ is an integral

$$f(x, \eta)=\int_{a'}^{\eta} F(x, y)dy,$$

and then

$$\psi(\eta)=\int_{a}^{\infty} dx \int_{a'}^{\eta} F(x, y)dy,$$

so that

$$\mathop{\mathcal{L}}_{\eta\to\infty} \psi(\eta)=\mathop{\mathcal{L}}_{\eta\to\infty}\int_{a}^{\infty} dx \int_{a'}^{\eta} F(x, y)dy=\int_{a}^{\infty} dx \int_{a'}^{\infty} F(x, y)dy,$$

since

$$\mathop{\mathcal{L}}_{\eta\to\infty}\int_{a}^{\infty} dx \int_{a'}^{\eta} F(x, y)dx=\int_{a}^{\infty} dx \left[\mathop{\mathcal{L}}_{\eta\to\infty}\int_{a'}^{\eta} F(x, y)dy\right].$$

Two useful theorems have been given by Bromwich (*Infinite Series*, 2nd Ed. p. 485 (4), and p. 490) which may be derived by a slight change in the proof of Theorem III. In both cases the upper limit of the integral $\psi(\eta)$ is, not ∞, but N, where N is a function of η, $N(\eta)$ say, that tends to ∞ when $\eta\to\infty$.

THEOREM IV. *Analogue of Tannery's Theorem. Let* $\psi(\eta)=\int_{a}^{N} f(x, \eta)dx$. *Condition* (ii) *of Theorem III remains, but, instead of condition* (i), *it is given that* $|f(x, \eta)|<M(x)$ *if* $\eta \geqq a'$, *where* $\int_{a}^{\infty} M(x)dx$ *is convergent. The theorem is then*

$$\mathop{\mathcal{L}}_{\eta\to\infty}\int_{a}^{N} f(x, \eta)dx=\int_{a}^{\infty}\varphi(x)dx.$$

THEOREM V. *Let* $\psi(\eta)=\int_{a}^{N} f(x,\eta)g(x)dx$. *Condition* (ii) *of Theorem III remains, but, instead of condition* (i), *it is given* (α') *that, when* η *is constant,* $f(x, \eta)$ *is a positive, monotonic, decreasing function of* x *such that* $f(a, \eta)<K$, *a constant, for every value of* η, *and* (β') *that the integral* $\int_{a}^{\infty} g(x)dx$ *is convergent. The theorem is then*

$$\mathop{\mathcal{L}}_{\eta\to\infty}\int_{a}^{N} f(x,\eta)g(x)dx=\int_{a}^{\infty}\varphi(x)g(x)dx.$$

The only difference from the proof of Theorem III is in the grounds on which $|\mu|$ and $|\nu|$ in equation (1) and $|\beta|$ in equation (2) can be made less than ε by choice of b. The

number N may always be taken larger than any chosen value of b. Now, taking the integral μ as typical, we have for Theorem IV,

$$\mu = \left|\int_b^N f(x, \eta')\,dx\right| < \int_b^N M(x)\,dx,$$

and this integral may, by choice of b, be made less than ε since the integral of $M(x)$ converges at ∞. Again, for Theorem V,

$$\mu = \left|\int_b^N f(x, \eta')g(x)\,dx\right| = \left|f(b, \eta')\int_b^\xi g(x)\,dx\right| < K\left|\int_b^\xi g(x)\,dx\right|,$$

where $\xi > b$; as before, b may be chosen so that $|\mu| < \varepsilon$, since $\int_a^\infty g(x)\,dx$ converges.

Ex. 1. If $\lambda > 0$ and if the integral $\int_0^\infty \varphi(x)\,dx$ converges, prove that

$$\underset{\lambda \to 0}{\mathcal{L}} \int_0^\infty e^{-\lambda x}\varphi(x)\,dx = \int_0^\infty \varphi(x)\,dx.$$

Apply Abel's Test and Theorem I; for detailed proof see *E.T.* p. 463.

Ex. 2. Show that (i) $\underset{a \to 0}{\mathcal{L}} \int_0^\infty \frac{\sin x\,dx}{\sqrt{(x^2 - a^2)}} = \frac{\pi}{2}$;

$$\text{(ii)}\ \underset{a \to 0}{\mathcal{L}} \int_0^\infty e^{-x^2 - \frac{a^2}{x^2}}\,dx = \int_0^\infty e^{-x^2}\,dx = \tfrac{1}{2}\sqrt{\pi}.$$

Ex. 3. If $f(x)$ is a positive, monotonic, decreasing (non-increasing) function of x for the range $x \geqq 0$, prove that

$$\underset{\lambda \to \infty}{\mathcal{L}} \int_0^\infty f(x)\frac{\sin\lambda(x-c)}{x-c}\,dx = \frac{\pi}{2}\{f(c+0) + f(c-0)\}, \quad c > 0,$$

$$= \frac{\pi}{2}f(+0), \qquad c = 0,$$

$$= 0, \qquad c < 0,$$

where $f(c+0)$ and $f(c-0)$ are the limits of $f(c+x)$ and $f(c-x)$ respectively for $x \to 0$ while $f(+0)$ is the limit of $f(x)$ for $x \to 0$, x being in all cases positive. (These limits exist since $f(x)$ is monotonic.)

(i) Take $b > |c|$; then, by the Second Theorem of Mean Value,

$$\int_b^\infty f(x)\frac{\sin\lambda(x-c)}{x-c}\,dx = f(b)\int_b^\xi \frac{\sin\lambda(x-c)}{x-c}\,dx = f(b)\int_{\lambda(b-c)}^{\lambda(\xi-c)} \frac{\sin y}{y}\,dy, \quad \text{...(1)}$$

where $\lambda(x-c) = y$ and $\xi \geqq b$. Since the integral of $\sin y/y$ converges at ∞, the above integral tends to zero when $\lambda \to \infty$ whether c be positive, zero, or negative.

(ii) If $0 < c < b$, we have

$$\int_0^b f(x)\frac{\sin\lambda(x-c)}{x-c}dx = \int_0^c f(x)\frac{\sin\lambda(x-c)}{x-c}dx + \int_c^b f(x)\frac{\sin\lambda(x-c)}{x-c}dx \quad \ldots(2)$$

$$= \int_0^{\lambda c} f\left(c-\frac{y}{\lambda}\right)\frac{\sin y}{y}dy + \int_0^{\lambda(b-c)} f\left(c+\frac{y}{\lambda}\right)\frac{\sin y}{y}dy$$

by putting $-y$ for $\lambda(x-c)$ in the first integral and $+y$ for $\lambda(x-c)$ in the second.

By Theorem IV we now find that, when $\lambda \to \infty$,

$$\mathop{\mathcal{L}}_{\lambda\to\infty}\int_0^b f(x)\frac{\sin\lambda(x-c)}{x-c}dx = f(c-0)\int_0^\infty \frac{\sin y}{y}dy + f(c+0)\int_0^\infty \frac{\sin y}{y}dy$$

$$= \frac{\pi}{2}\{f(c-0)+f(c+0)\}.$$

(iii) If $0 = c < b$ the first of the two integrals in the second member of equation (2) disappears and the limit is $\frac{1}{2}\pi(f+0)$.

The results in (i), (ii) and (iii) prove the theorem when $c \geqq 0$.

(iv) The result in (i) holds if $c < 0$, say $c = -c'$ where $c' > 0$. For

$$\int_0^b f(x)\frac{\sin\lambda(x+c')}{x+c'}dx = f(0)\int_0^\xi \frac{\sin\lambda(x+c')}{x+c'}dx = f(0)\int_{\lambda c'}^{\lambda(\xi+c')}\frac{\sin y}{y}dy$$

by putting y for $\lambda(x+c')$. Hence, since $c' > 0$, the integral tends to zero when $\lambda \to \infty$ and therefore the theorem is true when $c < 0$.

Of course, if $f(x)$ is continuous, $f(c+0) = f(c-0) = f(c)$.

Ex. 4. The theorem of Ex. 3 holds if $f(x)$ is bounded, but is a monotonic, increasing (non-decreasing) function for the range $0 \leqq x \leqq a$, and a monotonic, decreasing (non-increasing) function for the range $x \geqq a$; more generally if $f(x)$ is bounded and has only a finite number of turning values.

Since $f(x)$ is bounded, the positive constant A may be chosen so that $A + f(x)$ is positive for $x \geqq 0$. Now, when $f(a)$ is the only turning value of $f(x)$ and $f(x) < f(a)$ for $x < a$, let $F(x)$ and $G(x)$ be functions defined by the equations

$$\left.\begin{aligned} F(x) &= A + f(a) \\ G(x) &= A + f(a) - f(x) \end{aligned}\right\} 0 \leqq x \leqq a; \quad \left.\begin{aligned} F(x) &= A + f(x) \\ G(x) &= A \end{aligned}\right\} x \geqq a.$$

Each of the functions $F(x)$ and $G(x)$ is positive, monotonic and non-increasing for the range $x \geqq 0$, and therefore the theorem of Ex. 3 holds for each and therefore also for the difference $F(x) - G(x)$, which is $f(x)$.

It is easy to adapt the proof to the more general case of a finite number of turning values, but the above case is sufficient for the special application to be made of the theorem (see § 163, Exs. 1, 2).

160. Repeated Integrals: One Limit Infinite. In this article the following notations will be used:

$$f(y) = \int_a^x F(x, y)dx, \quad g(x) = \int_{a'}^{b'} F(x, y)dy.$$

The discontinuities of $F(x, y)$ satisfy the conditions of § 154.

THEOREM I. *If the integral $f(y)$ converges uniformly for the range $a' \leqq y \leqq b'$, and if the integral $g(x)$ converges uniformly for the arbitrarily large range $a \leqq x \leqq b$, then, η being any* fixed *number in (a', b'), the integrals*

$$\int_{a'}^{\eta} dy \int_a^{\infty} F(x, y)\,dx \;\ldots\;(1) \quad\text{and}\quad \int_a^{\infty} dx \int_{a'}^{\eta} F(x, y)\,dy \;\ldots\ldots\ldots(2)$$

exist and are equal.

By Theorem I of § 159, $f(y)$ is a continuous function of y for the range $a' \leqq y \leqq b'$, and therefore the integral (1) exists.

Again, by Theorem I of § 154, if η is any fixed number in (a', b') and b any fixed number greater than a,

$$\int_a^b dx \int_{a'}^{\eta} F(x, y)\,dy = \int_{a'}^{\eta} dy \int_a^b F(x, y)\,dx$$

$$= \int_{a'}^{\eta} dy \int_a^{\infty} F(x, y)\,dx - \int_{a'}^{\eta} dy \int_b^{\infty} F(x, y)\,dx.$$

Now, by the uniform convergence of $f(y)$, the number B can be chosen so that, if $b \geqq B$ and $a' \leqq y \leqq b'$, ε being as usual,

$$\left|\int_b^{\infty} F(x, y)\,dx\right| < \varepsilon, \quad \left|\int_{a'}^{\eta} dy \int_b^{\infty} F(x, y)\,dx\right| < (\eta - a')\varepsilon \leqq (b' - a')\varepsilon,$$

and therefore this repeated integral tends to zero when $b \to \infty$. Hence

$$\int_a^{\infty} dx \int_{a'}^{\eta} F(x, y)\,dy = \mathop{\mathcal{L}}_{b \to \infty} \int_a^b dx \int_{a'}^{\eta} F(x, y)\,dy = \int_{a'}^{\eta} dy \int_a^{\infty} F(x, y)\,dx,$$

so that the integral (2) exists and is equal to the integral (1).

Cor. The integral $\displaystyle\int_a^{\infty} dx \int_{a'}^{\eta} F(x, y)\,dy$

converges uniformly for the range $a' \leqq \eta \leqq b'$ when the integrals $f(y)$ and $g(x)$ satisfy the conditions of Theorem I; because the integral (1) is a continuous function of η for that range and therefore also the integral (2).

Ex. 1. Show that the integral $f(y)$ of Ex. 7, § 158, converges uniformly in the interval $(1, b')$, where b' is arbitrarily large.

De la Vallée Poussin observes that the uniform convergence of a

repeated integral may often be established by applying the corollary just stated. In the example, the integral

$$\int_0^\infty \Gamma(t)\{e^{-x}-(1+x)^{-t}\}\frac{dx}{x}$$

converges uniformly, as is easily proved, for the range $1 \leqq t \leqq b'$, where b' is arbitrarily large, while the integral

$$\int_1^{b'} \frac{\Gamma(t)}{x}\{e^{-x}-(1+x)^{-t}\}dt$$

converges uniformly in the arbitrary interval $(0, b)$. (The value of $\{e^{-x}-(1+x)^{-t}\}x^{-1}$ for $x=0$ may be taken to be $(t-1)$, which is the limit of this expression when $x \to 0$.) Hence, by the corollary, the repeated integral

$$\int_0^\infty \frac{dx}{x}\int_1^y \Gamma(t)\{e^{-x}-(1+x)^{-t}\}dt$$

converges uniformly for the range $1 \leqq y \leqq b'$.

Ex. 2. If $u(\xi)=\int_{a'}^{b'}dy\int_a^\xi F(x,y)dx$ and if $u(\xi)$ converges uniformly for the *unlimited* range $\xi \geqq a$, show that

$$\mathop{\mathcal{L}}_{N\to\infty}\int_{a'}^{b'}dy\int_N^\infty F(x,y)dx=0.$$

The condition that $u(\xi)$ should tend to a limit when $\xi \to \infty$ is that there should be a number N such that

$$|u(\xi_2)-u(\xi_1)|<\varepsilon \text{ if } \xi_2>\xi_1 \geqq N.$$

Now $u(\xi)$ converges uniformly through the unlimited range $\xi \geqq a$ and therefore $u(\infty)$ is a definite number. Hence

$$|u(\infty)-u(\xi_1)| \leqq \varepsilon \text{ if } \xi_1 \geqq N \text{ or } \mathop{\mathcal{L}}_{N\to\infty}\{u(\infty)-u(N)\}=0.$$

But $$u(\infty)-u(N)=\int_{a'}^{b'}dy\int_N^\infty F(x,y)dx.$$

Ex. 3. Show that $u(\xi)$ is a continuous function of ξ in the arbitrarily large interval (a, b).

THEOREM II. *The integrals* (1) *and* (2) *of Theorem I exist and are equal when the following conditions are satisfied:*

(i) *The integral $f(y)$ is only uniformly convergent in general for the range $a' \leqq y \leqq b'$;*

(ii) *The integral $g(x)$ converges uniformly for the arbitrarily large range $a \leqq x \leqq b$;*

(iii) *The integral $v(\eta_1)$, where*

$$v(\eta_1)=\int_a^\infty dx\int_{a'}^{\eta_1} F(x,y)dy,$$

converges uniformly for the range $a' \leqq \eta_1 \leqq b'$. (The symbol η_1 is used because η in the integral (2) is supposed to be fixed.)

We may suppose that $f(y)$ ceases to converge uniformly for only one value c' of y, where $a' < c' < \eta \leqq b'$; then $f(y)$ converges uniformly in $(a', c'-\delta)$ and $(c'+\delta', \eta)$, where δ and δ' have the usual meaning, and therefore, by Theorem I,

$$\int_{a'}^{c'-\delta} dy \int_a^\infty F(x, y)\,dx + \int_{c'+\delta'}^{\eta} dy \int_a^\infty F(x, y)\,dx$$
$$= \int_a^\infty dx \int_{a'}^{c'-\delta} F(x, y)\,dy + \int_a^\infty dx \int_{c'+\delta'}^{\eta} F(x, y)\,dy.$$

Now, by § 159, Theorem II, $v(\eta_1)$ is a continuous function of η_1 and therefore the limits for δ and δ' tending to zero of the integrals in the second member of this equation are obtained by making δ and δ' zero. Hence

$$\mathop{\mathcal{L}}_{\delta \to 0} \int_{a'}^{c'-\delta} dy \int_a^\infty F(x, y)\,dx + \mathop{\mathcal{L}}_{\delta' \to 0} \int_{c'+\delta'}^{\eta} dy \int_a^\infty F(x, y)\,dx$$
$$= \int_a^\infty dx \int_{a'}^{c'} F(x, y)\,dy + \int_a^\infty dx \int_{c'}^{\eta} F(x, y)\,dy;$$

that is, $$\int_{a'}^{\eta} dy \int_a^\infty F(x, y)\,dx = \int_a^\infty dx \int_{a'}^{\eta} F(x, y)\,dy,$$

by combining the respective pairs of integrals.

Cor. If $F(x, y)$ does not change sign, $v(\eta_1)$ will be a continuous function of η_1, provided the integral (2) exists. Therefore, when conditions (i) and (ii) are satisfied, the integral (1) will converge and be equal to the integral (2), when that integral converges.

Ex. 4. $$\int_0^\lambda dy \int_a^\infty e^{-x^2 y}\,dx = \int_a^\infty \frac{1 - e^{-\lambda x^2}}{x^2}\,dx, \ \lambda > 0,\ a \geqq 0.$$

Here $$\int_b^\infty e^{-x^2 y}\,dx = \frac{1}{\sqrt{y}} \int_{b\sqrt{y}}^\infty e^{-u^2}\,du,$$

so that the integral $f(y)$ only converges uniformly in general in $(0, \lambda)$; it converges uniformly for the range $0 < \delta \leqq y \leqq \lambda$.

Next $$g(x) = \int_0^\lambda e^{-x^2 y}\,dy = \frac{1 - e^{-\lambda x^2}}{x^2}.$$

When $x \to 0$, $g(x) \to \lambda$ and λ will be taken as the value of $g(x)$ when $x = 0$; the integral is obviously continuous in the arbitrarily large interval $(0, b)$.

The function $e^{-x^2 y}$ is always positive and the integral $v(\lambda)$ is

$$\int_a^\infty \frac{1 - e^{-\lambda x^2}}{x^2}\,dx,$$

which is manifestly convergent, so that the corollary is applicable.

THEOREM III. *If the integrals* $f(y)$ *and* $g(x)$ *are each only uniformly convergent in general in the respective intervals* (a', b') *and* (a, b), *the integrals* (1) *and* (2) *of Theorem I will exist and be equal if the integral* $u(\xi)$, *where*

$$u(\xi)=\int_{a'}^{b'} dy \int_a^{\xi} F(x, y)\,dx,$$

converges uniformly for the unlimited *range* $\xi \geqq a$.

By the conditions $u(\xi)$ converges when $\xi \to \infty$; that is, the integral $u(\infty)$, and therefore also the integral (1), exists.

Again, the conditions of Theorem IV, § 154, are satisfied, since $u(\xi)$ converges uniformly in (a, b), however large b may be, and therefore

$$\int_a^b dx \int_{a'}^{\eta} F(x, y)\,dy=\int_{a'}^{\eta} dy \int_a^b F(x, y)\,dx.$$

Now, by Ex. 2,

$$\mathop{\mathcal{L}}_{q\to\infty}\int_{a'}^{\eta} dy \int_a^b F(x, y)\,dx=\int_{a'}^{\eta} dy \int_a^{\infty} F(x, y)\,dx-\mathop{\mathcal{L}}_{b\to\infty}\int_{a'}^{\eta} dy \int_b^{\infty} F(x, y)\,dx$$
$$=\int_{a'}^{\eta} dy \int_a^{\infty} F(x, y)\,dx,$$

so that
$$\mathop{\mathcal{L}}_{b\to\infty}\int_a^b dx \int_{a'}^{\eta} F(x, y)\,dy=\int_{a'}^{\eta} dy \int_a^{\infty} F(x, y)\,dx,$$

and therefore the integral (2) exists and is equal to the integral (1).

Of course the theorem holds if one of the integrals $f(y)$ and $g(x)$ converges uniformly and the other only converges uniformly in general.

Cor. If $F(x, y)$ does not change sign $u(\xi)$ will be a continuous function of ξ and therefore, when $f(y)$ and $g(x)$ are only uniformly convergent in general, the integral (2) will exist and be equal to the integral (1) when that integral exists.

Ex. 5. Evaluate $\displaystyle\int_0^{\infty} \frac{dx}{(1+x)^{m+1}x^n}\int_0^1 y^{l-1}(1+xy)^m\,dy,$

when l, m, n are such that the integral converges.

If the integrals (1) and (2) of Theorem I are denoted by I_1 and I_2 respectively, this integral is I_2. The integrand does not change sign, so that we first test the convergence of the integrals $f(y)$ and $g(x)$; if these satisfy the conditions of Theorem III we next examine whether the integral I_1 converges. If I_1 converges so does I_2, and $I_2=I_1$.

Suppose, to begin with, that $m \geqq 0$. Then $(1+xy)/(1+x)$ lies between 1 and y and $[(1+xy)/(1+x)]^m$ lies between 1 and y^m, so that it is positive and not greater than unity.

$$(1)\quad f(y)=y^{l-1}\int_0^\infty \left(\frac{1+xy}{1+x}\right)^m \frac{dx}{(1+x)x^n} < y^{l-1}\int_0^\infty \frac{dx}{(1+x)x^n}.$$

For convergence at 0 we must have $n<1$, and for convergence at ∞, $n>0$; if $0<n<1$ the integral $f(y)$ converges when y is not zero, or even when $y=0$ if $l \geqq 1$. Thus, if $l>0$, $f(y)$ converges uniformly in the range $0<\delta \leqq y \leqq 1$ and converges uniformly in general in $(0, 1)$.

$$(2)\quad g(x)=\frac{1}{x^n}\int_0^1 \frac{y^{l-1}}{1+x}\left(\frac{1+xy}{1+x}\right)^m dy < \frac{1}{(1+x)x^n}\int_0^1 y^{l-1}\,dy.$$

For convergence at 0 we must have $l>0$ and then, since $0<n<1$, the integral $g(x)$ converges uniformly for the unlimited range $x \geqq \delta > 0$, and converges uniformly in general in any arbitrarily large interval $(0, b)$.

$$(3)\quad I_1=\int_0^1 y^{l-1}\,dy\int_0^\infty \left(\frac{1+xy}{1+x}\right)^m \frac{dx}{(1+x)x^n}$$
$$=\int_0^1 y^{l-1}\,dy\int_0^b \left(\frac{1+xy}{1+x}\right)^m \frac{dx}{(1+x)x^n}+R_b$$

The repeated integral with finite limits is obviously convergent when $l>0$, $m \geqq 0$, $0<n<1$. Further

$$(4)\quad R_b=\int_0^1 y^{l-1}\,dy\int_b^\infty \left(\frac{1+xy}{1+x}\right)^n \frac{dx}{(1+x)x^n} < \int_0^1 y^{l-1}\,dy\int_b^\alpha \frac{dx}{(1+x)x^n},$$

and clearly $R_b \to 0$ when $b \to \infty$, so that I_1 is convergent.

Hence the integral I_2 also converges and $I_2=I_1$, so that the value of the repeated integral is not changed by changing the order of integration.

In the integral I_2 change the variable y to v by the substitution $y/(1+xy)=v/(1+x)$; then

$$_1=\int_0^\infty \frac{dx}{x^n}\int_0^1 \frac{v^{l-1}\,dv}{(1+x-xv)^{l+m+1}}.$$

Next change the order of integration, as we are entitled to do by what has been proved, and then substitute the variable u for the variable x, where $x(1-v)=u$. We thus find that

$$I_2=\int_0^1 v^{l-1}(1-v)^{n-1}\,dv\int_0^\infty \frac{u^{-n}\,du}{(1\pm u)^{l+m+1}}$$
$$=B(l, n)\ B(1-n, l+m+n).$$

(See *E.T.* p. 350, Ex. 20, for the value of the integral with ∞ as limit).

Hence $$I_2=\frac{\pi}{\sin n\pi}\frac{\Gamma(l)\,\Gamma(l+m+n)}{\Gamma(l+m+1)\,\Gamma(l+n)}.$$

This result suggests that the restriction $m \geqq 0$ is too narrow. If $m<0$, then

$$1<[(1+xy)/(1+x)]^m<y^m,$$

and in (1), (2), (4) the power of y that remains after the substitution of y^m for $[(1+xy)/(1+x)]^m$ is y^{l+m-1}, so that m may be negative if $l+m>0$—a condition that implies $l>0$.

161. Differentiation under the Integral Sign. Let $f(y)$ and $\varphi(y)$ denote the integrals

$$f(y)=\int_a^\infty F(x, y)\,dx, \quad \varphi(y)=\int_a^\infty \frac{\partial F(x, y)}{\partial y}\,dx.$$

If $F(x, y)$, $\partial F(x, y)/\partial y$ and the integral $\varphi(y)$ satisfy conditions (i), (ii) and (iii) respectively of § 155 when $b=\infty$—that is, when $x \geqq a$ and $a' \leqq y \leqq b'$—then $\varphi(y)$ is the derivative of $f(y)$. No change at all is needed in the proof of the theorem when b is finite; the change of order of integration in equation (3) of § 155 is valid in the present case by Theorem I of § 160.

It may happen that the integral $\varphi(y)$, obtained by differentiating under the integral sign, does not converge; see, for example, Ex. 4, p. 468, of the *Elementary Treatise*, where the difficulty is overcome by a special device. De la Vallée Poussin has given a general method of dealing with such cases that is frequently successful.

Suppose that $F(x, y)$ and $\partial F(x, y)/\partial y$ are continuous functions of x and y for the ranges $a \leqq x \leqq b$, where b is arbitrarily large, and $a' \leqq y \leqq b'$; then, taking for brevity the symbol y to denote both the variable of integration and the upper limit of the integral, we have

$$F(x, y)-F(x, a')=\int_{a'}^{y} \frac{\partial F(x, y)}{\partial y}\,dy,$$

and therefore

$$\int_a^b F(x, y)\,dx-\int_a^b F(x, a')\,dx=\int_a^b dx \int_{a'}^{y} \frac{\partial F(x, y)}{\partial y}\,dy.$$

Change the order of integration in the repeated integral, as it is legitimate to do since $\partial F/\partial y$ is continuous, and let b tend to infinity; then

$$f(y)-f(a')=\underset{b\to\infty}{\mathcal{L}} \int_{a'}^{y} dy \int_a^b \frac{\partial F(x, y)}{\partial y}\,dx, \quad\text{..............(1)}$$

and

$$\frac{df(y)}{dy}=\frac{d}{dy}\left[\underset{b\to\infty}{\mathcal{L}} \int_{a'}^{y} dy \int_a^b \frac{\partial F(x, y)}{\partial y}\,dx\right]. \quad\text{.......(2)}$$

Now it may be possible to obtain a transformation of the form

$$\int_a^b \frac{\partial F(x, y)}{\partial y}\,dx=\varphi(b, y)+\int_a^b \psi(x, y)\,dx,$$

where the functions $\varphi(b, y)$ and $\psi(x, y)$ are such that

$$\mathop{\mathcal{L}}_{b\to\infty}\int_{a'}^{y}\varphi(b, y)\,dy=0$$

and the integral $\int_a^\infty \psi(x, y)\,dx$ converges uniformly for the range $a'\leqq y\leqq b'$. If these conditions are satisfied,

$$\int_{a'}^{y}dy\int_a^b\frac{\partial F(x, y)}{\partial y}\,dx=\int_{a'}^{y}\varphi(b, y)\,dy+\int_{a'}^{y}dy\int_a^b\psi(x, y)\,dx,$$

and therefore, when $b\to\infty$, equation (1) gives

$$f(y)-f(a')=\int_{a'}^{y}dy\int_a^\infty\psi(x, y)\,dx-\mathop{\mathcal{L}}_{b\to\infty}\int_{a'}^{y}dy\int_b^\infty\psi(x, y)\,dx$$

$$=\int_{a'}^{y}dy\int_a^\infty\psi(x, y)\,dx,$$

since the integral of $\psi(x, y)$ converges uniformly. Equation (2) then gives

$$\frac{df(y)}{dy}=\int_a^\infty\psi(x, y)\,dx.$$

Ex. 1. If $u=\int_0^\infty\cos(x^3-xy)\,dx$, prove that

$$\frac{d^2u}{dy^2}+\tfrac{1}{3}yu=0.$$

By Ex. 8 of § 158 the first derivative of u is given by the equation

$$\frac{du}{dy}=\int_0^\infty x\sin(x^3-xy)\,dx,$$

since the integral converges uniformly in an arbitrarily large interval $(0, b')$; the integral obtained by the second differentiation is, however, not convergent. But

$$\frac{\partial}{\partial y}x\sin(x^3-xy)=-x^2\cos(x^3-xy)$$

and

$$\int_0^b-x^2\cos(x^3-xy)\,dx=-\tfrac{1}{3}\int_0^b(3x^2-y)\cos(x^3-xy)\,dx-\tfrac{1}{3}y\int_0^b\cos(x^3-xy)\,dx,$$

so that $\quad\varphi(b, y)=-\tfrac{1}{3}\sin(b^3-by),\ \ \psi(x, y)=-\tfrac{1}{3}y\cos(x^3-xy).$

Now $$\mathop{\mathcal{L}}_{b\to\infty}\int_0^y\varphi(b, y)\,dy=\mathop{\mathcal{L}}_{b\to\infty}\frac{\cos(b^3-by)-\cos(b^3)}{3b}=0,$$

and the integral of $\psi(x, y)$ converges uniformly in $(0, b')$.

For $$y\int_B^\infty\cos(x^3-xy)\,dx=y\int_B^\infty\frac{3x^2-y}{3x^2}\cos(x^3-xy)\,dx+\frac{y^2}{3}\int_B^\infty\frac{\cos(x^3-xy)}{x^2}\,dx,$$

so that $$\left|y\int_B^\infty\cos(x^3-xy)\,dx\right|\leqq\frac{2b'}{3B^2}+\frac{b'^2}{3B}.$$

Hence, $$\frac{d^2u}{dy^2}=\int_0^\infty \psi(x, y)dx = -\tfrac{1}{3}y\int_0^\infty \cos(x^3 - xy)dx,$$

that is, $$\frac{d^2u}{dy^2}+\tfrac{1}{3}yu=0.$$

Ex. 2. If $u=\int_0^\infty \frac{\cos(xy)dx}{1+x^2}$, prove that $(y>0)$ $\frac{d^2u}{dy^2}=u$

The integral obtained by the first different ation converges uniformly; for the second derivative the integrand is $-x^2\cos(xy)/(1+x^2)$ and

$$\int_0^b \frac{-x^2\cos xy}{1+x^2}dx = -\frac{\sin by}{y}+\int_0^b \frac{\cos(xy)dx}{1+x^2}.$$

If $y\geqq c>0$, $$\lim_{b\to\infty}\int_c^y \frac{\sin by}{y}dy=0,$$

and the integral $\int_0^\infty \frac{\cos(xy)dx}{1+x^2}$ converges uniformly $(y\geqq c>0)$. Hence

$$\frac{d^2u}{dy^2}=\int_0^\infty \frac{\cos(xy)dx}{1+x^2}=u.$$

Deduce the results of Ex. 4, p. 468, of the *Elementary Treatise.*

Ex. 3. Let a and b be two numbers, real or complex, neither of them being zero; a and b, when real, are positive, but, when complex, have their real parts either positive or zero. If $y\geqq 0$, show that

$$u\equiv\int_0^\infty e^{-xy}(e^{-ax}-e^{-bx})\frac{dx}{x}=\int_y^\infty\left(\frac{1}{a+y}-\frac{1}{b+y}\right)dy.$$

Suppose $a=\alpha+i\alpha'$ and $b=\beta+i\beta'$; then

$$|e^{-ax}-e^{-bx}|\leqq e^{-\alpha x}+e^{-\beta x}.$$

The integrand of the integral u may be taken to be $(b-a)$ when $x=0$,

and $$\left|\int_B^\infty e^{-xy}(e^{-ax}-e^{-bx})\frac{dx}{x}\right|\leqq\int_B^\infty \frac{e^{-(\alpha+y)x}+e^{-(\beta+y)x}}{x}dx;$$

this integral tends to zero when $B\to\infty$ if $y\geqq 0$. [When $y=0$, $\alpha=0$, $\beta=0$, the integrals $\int_B^\infty \frac{e^{-i\alpha' x}}{x}dx$, $\int_B^\infty \frac{e^{-i\beta' x}}{x}dx$ both tend to zero when $B\to\infty$.]

Again, the integral obtained by differentiating u with respect to y is

$$-\int_0^\infty e^{-xy}(e^{-ax}-e^{-bx})dx=\frac{1}{b+y}-\frac{1}{a+y},$$

and this integral converges uniformly if $y\geqq a'>0$, so that it is equal to du/dy. Hence

$$\frac{du}{dy}=-\int_0^\infty e^{-xy}(e^{-ax}-e^{-bx})dx=\frac{1}{b+y}-\frac{1}{a+y}.$$

Now $u\to 0$ when $y\to\infty$, and therefore

$$u=\int_y^\infty\left(\frac{1}{a+y}-\frac{1}{b+y}\right)dy.$$

If $y \to 0$, we find (§ 159, Ex. 1, or *E.T.*, p. 463, Th. II)

$$\int_0^\infty \frac{e^{-ax}-e^{-bx}}{x}\,dx = \int_0^\infty \left(\frac{1}{a+y}-\frac{1}{b+y}\right)dy.$$

This result contains many particular cases, *e.g.* Ex. 26 (i) of § 146.

Again, let $a=1$, $b=i$; then

$$\int\left(\frac{1}{1+y}-\frac{1}{i+y}\right)dy = \int\left(\frac{1}{1+y}-\frac{y-i}{y^2+1}\right)dy$$

$$= \log\left\{\frac{1+y}{\sqrt{(1+y^2)}}\right\} + i\tan^{-1}y,$$

so that $\displaystyle\int_0^\infty \frac{(e^{-x}-\cos x)+i\sin x}{x}\,dx = i\frac{\pi}{2}$,

and therefore $\displaystyle\int_0^\infty \frac{e^{-x}-\cos x}{x}\,dx = 0,\ \int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$

EXERCISES XIX.

1. If $0 < a < b$, prove that

$$\int_0^\infty dx \int_a^b e^{-xy}\,dy = \int_a^b dy \int_0^\infty e^{-xy}\,dx,$$

and deduce that

$$\int_0^\infty \frac{e^{-ax}-e^{-bx}}{x}\,dx = \log\frac{b}{a}.$$

2. If $a>0$, $b>0$, prove that

$$\int_0^\infty dx \int_0^a e^{-(y-ib)x}\,dy = \int_0^a dy \int_0^\infty e^{-(y-ib)x}\,dx,$$

and deduce that

(i) $\displaystyle\int_0^\infty \frac{1-e^{-ax}}{x}e^{ibx}\,dx = \tfrac{1}{2}\log\left(1+\frac{a^2}{b^2}\right)+i\tan^{-1}\left(\frac{a}{b}\right);$

(ii) $\displaystyle\int_0^\infty \frac{1-e^{-ax}}{x}\cos bx\,dx = \tfrac{1}{2}\log\left(1+\frac{a^2}{b^2}\right);$

(iii) $\displaystyle\int_0^\infty \frac{1-e^{-ax}}{x}\sin bx\,dx = \tan^{-1}\left(\frac{a}{b}\right).$

3. If $a>0$ and $b \geqq 0$, show that

$$\int_0^\infty \frac{\sin bx}{x(a^2+x^2)}\,dx = \frac{\pi}{2a^2}(1-e^{-ab}).$$

4. If $a>0$ and $b>1$, prove that

(i) $\displaystyle\int_0^\infty \frac{\cos x}{x}\,dx \int_0^1 e^{-ay}\sin(xy)\,dy = \int_0^1 e^{-ay}\,dy \int_0^\infty \frac{\cos x\sin(xy)}{x}\,dx;$

(ii) $\displaystyle\int_0^\infty \frac{\cos x}{x}\,dx \int_1^b e^{-ay}\sin(xy)\,dy = \int_1^b e^{-ay}\,dy \int_0^\infty \frac{\cos x\sin(xy)}{x}\,dx.$

Show that each of the integrals in (i) is zero, and that each of the integrals in (ii) is equal to $\pi(e^{-a}-e^{-ab})/2a$.

5. Let $f(x)=Ae^{-ax}+Be^{-bx}+\ldots+Ke^{-kx}=\Sigma Ae^{-ax}$, where $a, b, \ldots k$ are either real and positive (not zero) or else complex, with real parts that are positive or zero, and $A, B, \ldots K$ constants. If u denotes the integral

$$\int_0^\infty e^{-xy}f(x)\frac{dx}{x},\ y>0,$$

prove that the integral is convergent if $\Sigma A=0$, and then show that

$$\int_0^\infty \{\Sigma Ae^{-ax}\}\frac{dx}{x}=-\Sigma A\log a.$$

6. If $\Sigma A=0$ and $\Sigma Aa=0$, the notation being as in Ex. 5, prove that

$$\int_0^\infty \{\Sigma Ae^{-ax}\}\frac{dx}{x^2}=\Sigma Aa\log a.$$

7. If $\Sigma A=0$, $\Sigma Aa=0$ and $\Sigma Aa^2=0$ (notation of Ex. 5), prove that

$$\int_0^\infty \{\Sigma Ae^{-ax}\}\frac{dx}{x^3}=-\tfrac{1}{2}\Sigma Aa^2\log a.$$

8. If $\Sigma A=0$ and $\Sigma Aa^r=0$, $r=1, 2, \ldots n-1$ (notation of Ex. 5), prove that

$$\int_0^\infty \{\Sigma Ae^{-ax}\}\frac{dx}{x^n}=\frac{(-1)^n}{(n-1)!}\Sigma Aa^{n-1}\log a.$$

9. In Ex. 5 suppose that $a, b, \ldots k$ are pure imaginary numbers, $a=\alpha i$, $b=\beta i$, $\ldots k=\kappa i$, $\alpha, \beta, \ldots \kappa$, positive, so that $\log a=\log\alpha+i\pi/2, \ldots \log k=\log\kappa+i\pi/2$. Deduce from Exs. 5, 6, 7 that, if the conditions that connect the constants $A, B, \ldots K$ in the respective examples are satisfied,

$$\text{(i)}\quad \int_0^\infty \Sigma(A\cos\alpha x)\,.\,\frac{dx}{x}=-\Sigma A\log\alpha;$$

$$\text{(ii)}\quad \int_0^\infty \Sigma(A\sin\alpha x)\frac{dx}{x^2}=-\Sigma A\alpha\log\alpha;$$

$$\text{(iii)}\quad \int_0^\infty \Sigma(A\cos\alpha x)\frac{dx}{x^3}=\tfrac{1}{2}\Sigma A\alpha^2\log\alpha.$$

10. Deduce from Ex. 8 the formulae corresponding to those in Ex. 9, the conditions that connect the constants $A, B, \ldots K$ being satisfied.

If n is odd, $n=2m+1$,

$$\int_0^\infty \Sigma(A\cos\alpha x)\,.\,\frac{dx}{x^{2m+1}}=\frac{(-1)^{m-1}}{(2m)!}\Sigma A\alpha^{2m}\log\alpha,$$

while if n is even, $n=2m$,

$$\int_0^\infty \Sigma(A\sin\alpha x)\,.\,\frac{dx}{x^{2m}}=\frac{(-1)^m}{(2m-1)!}\Sigma A\alpha^{2m-1}\log\alpha.$$

For other formulae of a like kind see the article by Hardy, quoted in § 146 (Frullani's Integral).

11. If a and b are either real and positive or else complex, with real parts that are positive or zero, show that the integral

$$\int_0^\infty \{e^{-ax}(1+Ax)-e^{-bx}(1+Bx)\}\frac{dx}{x^2}$$

is convergent if $A-a=B-b=c$, say; then prove that when $A=a+c$ and $B=b+c$ the integral is equal to

$$b-a+c\log(b/a).$$

Deduce, by taking $a=1$, $b=-i$, $c=i$, that

(i) $$\int_0^\infty \{e^{-x}(1+x)-\cos x\}\frac{dx}{x^2}=\frac{\pi}{2}-1,\quad \int_0^\infty (xe^{-x}-\sin x)\frac{dx}{x^2}=-1;$$

and, by taking $b=1$, $c=-(a+\frac{1}{2})$, that

(ii) $$\int_0^\infty \left\{(a-1)e^{-x}+\left(\frac{1}{x}-\frac{1}{2}\right)(e^{-ax}-e^{-x})\right\}\frac{dx}{x}=(a+\tfrac{1}{2})\log a-(a-1).$$

(Bromwich, *Inf. Ser.* 2nd Ed. p. 488.)

12. The integral $\displaystyle\int_0^\infty \{e^{-ax}-(A+Bx+Cx^2)e^{-bx}\}\frac{dx}{x^3}$

converges if $A=1$, $B=b-a$, $C=\frac{1}{2}(b-a)^2$, and its value is then

$$\tfrac{3}{4}a^2-ab+\tfrac{1}{4}b^2+\tfrac{1}{2}a^2\log\frac{b}{a}.$$

13. Deduce from Ex. 6, or prove independently, that

$$\int_0^1 \frac{(b-c)x^a+(c-a)x^b+(a-b)x^c}{(\log x)^2}\frac{dx}{x}=\Sigma(b-c)\,a\log a.$$

14. If $0<p<1$, $\displaystyle\int_0^1 \frac{x^{p-1}-x^{-p}}{\log x}\frac{dx}{1+x}=\log\tan\frac{p\pi}{2}.$

15. If $0<p<1$, show that

(i) $$\int_0^\infty \frac{x^{p-1}\,dx}{(1+x)^{n+1}}=\frac{(1-p)(2-p)\dots(n-p)}{n!}\frac{\pi}{\sin p\pi};$$

(ii) $$\int_0^\infty \frac{x^{p+n}\,dx}{(1+x)^{2n+2}}=(-1)^n\frac{p(p^2-1^2)(p^2-2^2)\dots(p^2-n^2)}{(2n+1)!}\frac{\pi}{\sin p\pi}$$

(Bertrand.)

16. Prove that if $a>0$,

(i) $$\int_0^\infty e^{-a^2x^2}\cosh 2bx\,dx=\frac{\sqrt{\pi}}{2a}e^{\frac{b^2}{a^2}};$$

(ii) $$\int_0^\infty e^{-a^2x^2}\sinh 2bx\,dx=\frac{1}{a}e^{\frac{b^2}{a^2}}\int_0^{\frac{b}{a}} e^{-x^2}\,dx.$$

17. Prove that

$$\int_0^\infty e^{-x^2}(x\sin 2x-\sin^2 x)\frac{dx}{x^2}=(1-e^{-1})\frac{\sqrt{\pi}}{2}.$$

(Bertrand.)

18. $\displaystyle\int_0^\infty \log\left(\frac{e^x+e^{-x}+2\sin\theta}{e^x+e^{-x}}\right)dx=\tfrac{1}{2}(\pi\theta-\theta^2),\ 0\leqq\theta\leqq\frac{\pi}{2}.$

19. If m is a positive integer and $b>0$,

$$\int_0^\infty \frac{\sin 2mbx}{\sin bx}\frac{dx}{1+x^2}=\frac{\pi(1-e^{-2mb})}{e^b-e^{-b}}.$$

20. If $m>0$ and $a>0$, prove that

$$\int_0^\infty \frac{\sin^2 mx\,dx}{x^2(x^2+a^2)}=\frac{\pi}{4a^3}(2ma-1+e^{-2ma}).$$

21. If $$u=\int_0^\infty \frac{\lambda x-\sin\lambda x}{x^3(x^2+a^2)}dx,\ \lambda>0,\ a>0, \text{ show that}$$

$$\frac{du}{d\lambda}=\frac{\pi}{2a^3}(\lambda a-1+e^{-\lambda a}),$$

and deduce that

$$\int^\infty x-\sin x\quad dx=\frac{\pi}{2a^4}(\tfrac{1}{2}a^2-a+1-e^{-a}).$$

22. If $P=\int_{-a} e^{-2at}\cos(t^2-a^2)dt,\ Q=\int_{-a}^\infty e^{-2at}\sin(t^2-a^2)dt,$ prove that $P+Q$ is independent of a, and state its value.

23. If $a>0$ and $b>0$, prove that

$$\int_0^\infty \tan^{-1}(ax)\tan^{-1}(bx)\frac{dx}{x^2}=\frac{\pi}{2}\log\left\{\frac{(a+b)^{a+b}}{a^a b^b}\right\}.$$

24. If $0\leqq\alpha<\frac{\pi}{2}$, show that

$$_{-\infty}\frac{\tan^{-1}x\,dx}{x^2-2x\sin\alpha+1}=\frac{}{2\cos\alpha}.$$

25. $$\int_0^\infty\left(\frac{\tan^{-1}x}{x}\right)^2dx=\frac{\pi}{2}\left(\log 8-\frac{\pi^2}{8}\right).$$

26. If $u=\int_0^\infty e^{-ax^2-bx^2}dx$, where $a>0$ and $b>0$, prove that,

$$3ab\frac{\partial^2u}{\partial b^2}-3a\frac{\partial u}{\partial b}-2b^2\frac{\partial u}{\partial a}=1.$$

27. If $$u_n=\int_0^{\frac{\pi}{2}}\frac{dx}{(a+b\tan^2x)^n}=\int_0^\infty\frac{dy}{(a+by^2)^n(1+y^2)},$$ where $a>0$ and $b>0$, prove that

(i) $$u_n=\frac{1}{n-1}\frac{du_{n-1}}{da}=\dots=\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}u_1}{da^{n-1}};$$

(ii) $$\frac{2}{\pi}u_n=\frac{1.3.5\dots(2n-3)}{2.4.6\dots(2n-2)}\frac{1}{a^{n-\frac{1}{2}}b^{\frac{1}{2}}}+\frac{(-1)^n}{(n-1)!}\frac{1}{b^{\frac{1}{2}}}\frac{d^{n-1}}{da^{n-1}}\left(\frac{1}{\sqrt{a}+\sqrt{b}}\right).$$

28. Show that, if $a>0$ and b is any real or complex number,

$$\int_{-\infty}^\infty e^{-ax^2-2bx}dx=\frac{\sqrt{\pi}}{\sqrt{a}}e^{\frac{b^2}{a}},$$

and deduce that $$\int e^{-ax^2}\cos 2bx\,dx=\frac{\sqrt{\pi}}{2\sqrt{a}}e^{-\frac{b^2}{a}}.$$

29. (i) If $u = a^2 \tan^2 x + b^2 \cot^2 x$ and $a > 0$, $b > 0$,

$$\int_0^{\frac{\pi}{2}} e^{-u} \sec^2 x \operatorname{cosec}^2 x \, dx = \frac{\sqrt{\pi}}{2} \frac{a+b}{ab} e^{-2ab};$$

(ii) If $v = \dfrac{\alpha^2}{x-a} + \dfrac{\beta^2}{b-x}$ where $b > a$, $\alpha > 0$, $\beta > 0$,

$$\int_a^b e^{-v} \frac{dx}{\{(x-a)(b-x)\}^{\frac{3}{2}}} = \sqrt{\pi} \frac{\alpha+\beta}{\alpha\beta(b-a)^{\frac{3}{2}}} e^{-\frac{(\alpha+\beta)^2}{b-a}}.$$

162. Repeated Integrals: Infinite Limits. The theorems of § 160 will now be extended to the case in which the upper limit of both integrals is infinite; the discontinuities of the integrand are understood to satisfy the conditions stated in § 154.

THEOREM I. *If the integrals*

$$\int_a^\infty F(x, y)\,dx \ldots (1), \quad \int_{a'}^\infty F(x, y)\,dy \quad \ldots\ldots\ldots\ldots\ldots(2)$$

converge uniformly through the arbitrary intervals (a', b') *and* (a, b) *respectively, and if the integral* $v(\eta)$, *where*

$$v(\eta) = \int_a^\infty dx \int_{a'}^\eta F(x, y)\,dy, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

converges uniformly for the unlimited *range* $\eta \geqq a'$, *then*

$$\int_{a'}^\infty dy \int_a^\infty F(x, y)\,dx = \int_a^\infty dx \int_{a'}^\infty F(x, y)\,dy. \quad \ldots\ldots\ldots\ldots(4)$$

Let the function $f(x, \eta)$ of Theorem III, § 159, be defined as

$$f(x, \eta) = \int_{a'}^\eta F(x, y)\,dy;$$

then the function $\psi(\eta)$ of that theorem is the integral $v(\eta)$. We now find

$$\begin{aligned}\int_{a'}^\infty dy \int_a^\infty F(x, y)\,dx &= \mathop{\mathcal{L}}_{\eta\to\infty} \int_{a'}^\eta dy \int_a^\infty F(x, y)\,dx \\ &= \mathop{\mathcal{L}}_{\eta\to\infty} \int_a^\infty dx \int_{a'}^\eta F(x, y)\,dy \quad \text{(Th. I, § 160)} \\ &= \int_a^\infty dx \int_{a'}^\infty F(x, y)\,dy. \quad \text{(Th. III, § 159)}\end{aligned}$$

Of course, if the integrals (1) and (2) converge uniformly for the *unlimited* ranges $y \geqq a'$ and $x \geqq a$ respectively, the theorem is true; a similar observation is applicable to the other theorems.

THEOREM II. *The equation* (4) *is true if the integral* (1) *is only uniformly convergent in general, provided the other conditions of Theorem I are satisfied.*

This result follows from Theorem II, § 160, in the same way as Theorem I follows from Theorem I, § 160.

THEOREM III. *If the integrals* (1) *and* (2) *are only uniformly convergent in general through the arbitrary intervals* (a', b') *and* (a, b) *respectively, but if the integral* $v(\eta)$ *and the integral* $u(\xi)$, *where*

$$u(\xi)=\int_{a'}^{\infty}dy\int_{a}^{\xi}F(x,\,y)\,dx \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3')$$

converge uniformly for the unlimited *ranges* $\eta \geqq a'$ *and* $\xi \geqq a$ *respectively, then equation* (4) *is true, provided one of the integrals in* (4) *is determinate.*

Suppose it is the integral in the second member of equation (4) that is determinate, and denote it by A; then the condition that the integral should converge is

$$\mathop{\mathcal{L}}_{b\to\infty}\int_{b}^{\infty}dx\int_{a'}^{\infty}F(x,\,y)\,dy=0. \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Again, since the integral $v(\eta)$ is continuous, we have, by Theorem II, § 160,

$$\int_{a'}^{\eta}dy\int_{a}^{\infty}F(x,\,y)\,dx=\int_{a}^{\infty}dx\int_{a'}^{\eta}F(x,\,y)\,dy=A-R(\eta), \;\ldots\ldots (6)$$

where $$R(\eta)=\int_{a}^{\infty}dx\int_{\eta}^{\infty}F(x,\,y)\,dy.$$

The theorem will therefore be proved if it is shown that $R(\eta)$ tends to zero when $\eta\to\infty$, since the integral in the first member of (6) becomes the other integral of equation (4) when $\eta\to\infty$.

Now $R(\eta)$ converges uniformly for $\eta \geqq a'$; for

$$\int_{b}^{\infty}dx\int_{\eta}^{\infty}F(x,\,y)\,dy=\int_{b}^{\infty}dx\int_{a'}^{\infty}F(x,\,y)\,dy-\int_{b}^{\infty}dx\int_{a'}^{\eta}F(x,\,y)\,dy$$
$$=\alpha-\beta, \text{ say.}$$

By (5) we can choose M' so that $|\alpha|<\varepsilon$ if $b>M'$ and, because the integral $v(\eta)$ converges uniformly for $\eta \geqq a'$, we can choose M'' so that $|\beta|<\varepsilon$ if $b>M''$. Let M be the greater of the numbers M' and M''; then $|\alpha-\beta|<2\varepsilon$ for $\eta \geqq a'$ if $b>M$, so that $R(\eta)$ converges uniformly for $\eta \geqq a'$.

Next we have

$$R(\eta)=\int_a^b dx\int_\eta^\infty F(x,\,y)\,dy+\int_b^\infty dx\int_\eta^\infty F(x,\,y)\,dy$$
$$=\int_\eta^\infty dy\int_a^b F(x,\,y)\,dx+\int_b^\infty dx\int_\eta^\infty F(x,\,y)\,dy$$

by Theorem III of § 160, since the integral $u(\xi)$ converges uniformly for $\xi \geqq a'$. Hence, if γ denote the first of these two integrals,

$$R(\eta)=\gamma+(\alpha-\beta).$$

If b is any fixed number greater than M (as determined above), $|\alpha-\beta|<2\varepsilon$. Further, the integral $u(\xi)$ converges uniformly for $\xi \geqq a$, and therefore we can choose N so that $|\gamma|<\varepsilon$ if $\eta>N$. Hence $|R(\eta)|<3\varepsilon$ if $\eta>N$, so that $R(\eta)\to 0$ when $\eta\to\infty$. The theorem is thus proved.

Cor. The theorems are considerably simplified if $F(x,\,y)$ does not change sign, because in that case $v(\eta)$ is continuous when $v(\infty)$ is determinate, and $u(\xi)$ is continuous when $u(\infty)$ is determinate. Theorems I and II are therefore true (when the integrals (1) and (2) converge as required), provided the integral $v(\infty)$ is determinate, and Theorem III, provided one of the integrals in (4) is determinate.

Ex. 1. Prove that, if $c>0$,

$$\text{(i)}\ \int_0^\infty e^{-cy}\,dy\int_0^\infty \frac{\cos x\sin(xy)}{x}\,dx=\int_0^\infty \frac{\cos x\,dx}{x^2+c^2}\,;$$

$$\text{(ii)}\ \int_0^\infty e^{-cy}\,dy\int_0^\infty \frac{\sin x\cos(xy)}{x}\,dx=\int_0^\infty \frac{\sin x}{x}\,\frac{c\,dx}{x^2+c^2}\,.$$

Let $F(x,\,y)=e^{-cy}\cos x\sin(xy)/x$ and

$$f(y)=\int_0^\infty F(x,y)\,dx,\quad g(x)=\int_0^\infty F(x,y)\,dy,\quad v(\eta)=\int_0^\infty dx\int_0^\eta F(x,y)\,dy.$$

By § 158, Ex. 2, $f(y)$ converges uniformly in general for $y \geqq 0$ and $g(x)$ converges uniformly for $x \geqq 0$, while, by § 158, Ex. 4, $v(\eta)$ converges uniformly for $\eta \geqq 0$. Hence, by Theorem II, the order of integration of the repeated integral in (i) may be changed, and the new integral is

$$\int_0^\infty \frac{\cos x}{x}\,dx\int_0^\infty e^{-cy}\sin(xy)\,dy=\int_0^\infty \frac{\cos x\,dx}{x^2+c^2}\,.$$

By using Ex. 3 and Ex. 5 of § 158 the second of the above equations may be proved in the same way.

Ex. 2. Deduce from Ex. 1 that, if $c>0$,

$$\text{(i)}\ \int_0^\infty \frac{\cos x\,dx}{x^2+c^2}=\frac{\pi}{2c}e^{-c};\quad \text{(ii)}\ \int_0^\infty \frac{\sin x\,dx}{x(x^2+c^2)}=\frac{\pi}{2c^2}(1-e^{-c}).$$

These results are obtained by evaluating the repeated integrals in Ex. 1, (i) and (ii) respectively. Note that

$$\cos x \sin(xy) = \tfrac{1}{2}\sin(y+1)x + \tfrac{1}{2}\sin(y-1)x,$$

and therefore

$$\int_0^\infty \frac{\cos x \sin(xy)}{x}\,dx = \tfrac{1}{2}\int_0^\infty \frac{\sin(y+1)x}{x}\,dx + \tfrac{1}{2}\int_0^\infty \frac{\sin(y-1)x}{x}\,dx$$

$$= \frac{1}{2}\cdot\frac{\pi}{2} + \frac{1}{2}\begin{cases} -\pi/2, & \text{if } 0 \leqq y < 1, \\ \pi/2, & \text{if } 1 < y; \end{cases}$$

so that

$$\int_0^\infty e^{-cy}\,dy \int_0^\infty \frac{\cos x \sin(xy)}{x}\,dx = \frac{\pi}{4c} - \frac{\pi}{4}\int_0^1 e^{-cy}\,dy + \frac{\pi}{4}\int_1^\infty e^{-cy}\,dy$$

$$= \frac{\pi}{4c} - \frac{\pi}{4}\left(\frac{1-e^{-c}}{c}\right) + \frac{\pi}{4c}e^{-c} = \frac{\pi}{2c}e^{-c}.$$

The repeated integral in (ii) gives in the same way

$$\frac{\pi}{2c}(1-e^{-c}).$$

Ex. 3. Prove that $\displaystyle\int_0^\infty \frac{e^{ix}}{\sqrt{x}}\,dx = (1+i)\sqrt{\left(\frac{\pi}{2}\right)}$.

This integral has been evaluated in the *Elementary Treatise* (p. 471, (9)); another method of evaluation will now be given as an illustration of Theorem III, and also as an example of the substitution of an integral for a given function (see *E.T.* p. 477, Ex. 2).

Denote the integral by w, and for $1/\sqrt{x}$ substitute the integral

$$\frac{2}{\sqrt{\pi}}\int_0^\infty e^{-xy^2}\,dy.$$

$$\left(\frac{\sqrt{\pi}}{2} = \int_0^\infty e^{-u^2}\,du = \sqrt{x}\int_0^\infty e^{-xy^2}\,dy, \quad \text{if } u = y\sqrt{x}.\right)$$

Hence

$$\frac{\sqrt{\pi}}{2}\cdot w = \int_0^\infty e^{ix}\,dx \int_0^\infty e^{-xy^2}\,dy \quad \text{.........................(i)}$$

$$= \int_0^\infty dy \int_0^\infty e^{-(y^2-i)x}\,dx, \quad \text{.......................(ii)}$$

provided the change of order of integration is legitimate. This change will be considered later. Now

$$\int_0^\infty e^{-(y^2-ix)}\,dx = \frac{1}{y^2-i} = \frac{y^2+i}{y^4+1},$$

so that

$$\frac{\sqrt{\pi}}{2}w = \int_0^\infty \frac{y^2\,dy}{y^4+1} + i\int_0^\infty \frac{dy}{y^4+1}.$$

Let $y^4 = u$, and it is easily seen that each of these integrals is equal to $\tfrac{1}{4}B(\tfrac{1}{4}, \tfrac{3}{4})$ (*E.T.* p. 350, Ex. 20, (ii)). But $\tfrac{1}{4}B(\tfrac{1}{4}, \tfrac{3}{4}) = \dfrac{\pi}{2\sqrt{2}}$, and therefore

$$w = (1+i)\sqrt{\left(\frac{\pi}{2}\right)},$$

so that, by equating real and imaginary parts, we find

$$\int_0^\infty \frac{\cos x}{\sqrt{x}}\,dx = \int_0^\infty \frac{\sin x}{\sqrt{x}}\,dx = \sqrt{\left(\frac{\pi}{2}\right)}.$$

We have now to consider the change of order of integration. The integral (ii) has been found to be determinate. The integrals

$$\int_0^\infty e^{ix}e^{-xy^2}\,dx \text{ and } \int_0^\infty e^{ix}e^{-xy^2}\,dy$$

converge uniformly for the unlimited ranges $y \geqq a' > 0$ and $x \geqq a > 0$ respectively. For

$$\left|\int_b^\infty e^{ix}e^{-xy^2}\,dx\right| \leqq \frac{1}{y^2}\int_{by^2}^\infty e^{-u}\,du, \quad \left|\int_{b'}^\infty e^{ix}e^{-xy^2}\,dy\right| \leqq \frac{1}{\sqrt{x}}\int_{b'\sqrt{x}}^\infty e^{-v^2}\,dv$$

by the substitutions $xy^2 = u$ and $y\sqrt{x} = v$, and the uniform convergence is obvious.

Let $v(\eta)$ and $u(\xi)$ be the integrals obtained by taking η as the upper limit of the y-integral in (i), and ξ as the upper limit of the x-integral in (ii). By § 158, Ex. 6, the integral $v(\eta)$ converges uniformly for the unlimited range $\eta \geqq 0$, and it is easily proved that the integral $u(\xi)$ converges uniformly for the unlimited range $\xi \geqq 0$. Hence all the conditions of Theorem III are satisfied, so that the change of order is legitimate.

Ex. 4. Prove that $B(m, n) = \Gamma(m)\Gamma(n)/\Gamma(m+n)$, $m > 0$, $n > 0$.

By the definition of $\Gamma(m+n)$ we have

$$\frac{\Gamma(m+n)}{(1+x)^{m+n}} = \int_0^\infty e^{-(1+x)y}\,y^{m+n-1}\,dy,$$

and therefore, multiplying by x^{n-1} and integrating with respect to x from 0 to ∞, we find (§ 145, Ex. 4, or *E.T.* p. 350, Ex. 20)

$$\Gamma(m+n)B(m, n) = \int_0^\infty x^{n-1}\,dx\int_0^\infty e^{-(1+x)y}\,y^{m+n-1}\,dy.$$

Change the order of integration and the repeated integral is simply $\Gamma(m)\Gamma(n)$, so that the equation is proved if the change of order is legitimate.

If $F(x, y) = e^{-(1+x)y}x^{n-1}y^{m+n-1}$ we have

$$\text{(i)}\quad \int_0^\infty F(x, y)dx = \Gamma(n)\,.\,e^{-y}y^{m-1};$$

$$\text{(ii)}\quad \int_0^\infty F(x, y)dy = \Gamma(m+n)x^{n-1}/(1+x)^{m+n}.$$

The integral (i) converges uniformly for the range $y \geqq a' > 0$, and the integral (ii) for the range $x \geqq a > 0$. Further, $F(x, y)$ is positive and the repeated integrals exist. Hence the change of order is legitimate.

Ex. 5. If $F(x, y) = e^{-x^2y}x\sin 2ax\sin y$, $a > 0$, show that

$$\int_0^\infty dy\int_0^\infty F(x, y)dx = \int_0^\infty dx\int_0^\infty F(x, y)dy.$$

Let $f(y) = \int_0^\infty F(x, y)dx = \sin y\int_0^\infty e^{-x^2y}x\sin 2ax\,dx$; then, if $y > 0$,

$$\int_b^\infty e^{-x^2y}x\sin 2ax\,dx = \frac{e^{-b^2y}\sin 2ab}{2y} + \frac{a}{y}\int_b^\infty e^{-x^2y}\cos 2ax\,dx,$$

and $\left|\int_b^\infty F(x, y)dx \right| \qquad \frac{\sin y}{y}\, e^{-b^2 y} + \frac{\sin y}{}\, e^{-b^2 y} < \frac{3}{2} e^{-b^2 y} < \frac{3}{2} e^{-b^2 a'}$

so that the integral $f(y)$ converges uniformly for $y \geqq a' > 0$.

Again, it is easy to show that $\int_0^\infty F(x, y)\, dy$ converges uniformly for $x \geqq 0$.

Now, let
$$v(\eta) = \int_0^\infty x \sin 2ax\, dx \int_0^\eta e^{-x^2 y} \sin y\, dy\,;$$
then
$$v(\eta) = \int_0^\infty \frac{(1 - e^{-\eta x^2} \cos \eta) x \sin 2ax}{x^4 + 1} dx - \sin \eta \int_0^\infty \frac{e^{-\eta x^2} x^3 \sin 2ax}{x^4 + 1} dx,$$
and $v(\eta)$ converges uniformly for $\eta \geqq 0$, as is very easily proved. Note that the integral
$$\int_0^\infty \frac{x^3 \sin 2ax}{x^4 + 1} dx$$
is convergent since, when x is large, $x^3/(x^4 + 1)$ is a positive, monotonic, decreasing function which tends to zero when x tends to infinity.

Thus, by Theorem II, the change of order is legitimate. Ex. 33, (i), (ii), (iii) on p. 482 of the *Elementary Treatise* may be taken in connection with this example.

163. Double Integrals with Infinite Limits. The evaluation of the double integral when the limits of the integral with respect to the variables (one or both) are infinite is usually effected by means of a repeated integral. It is possible that the double integral may exist, and yet not be equal to either of the repeated integrals; further, the two repeated integrals may exist and be equal and yet not be equal to the corresponding double integral. A detailed investigation of the matter is, however, outside the limits of this book, and the student is referred to the investigations by De la Vallée Poussin and Stolz (see § 156); he should also consult an article by Bromwich in the *Proceedings of the London Mathematical Society*, vol. i. (2nd Series), 1904, pp. 176-201, and his textbook on *Infinite Series* (2nd Ed.), pp. 503-513. Bromwich's definition admits conditionally convergent double integrals; we have followed the more usual practice of admitting only absolutely convergent double integrals.

It is hardly necessary to repeat the remarks made in § 156 on the evaluation of the (improper) double integral; the general procedure is the same in the present case as in the case there stated. The General Theorem quoted in § 156 also holds when the limits of the integral are infinite.

Again, as regards the change of variables reference may be made to what has been stated in § 157, with the examples there given.

Ex. 1. Prove that, if $a>0$ and $n \geqq 1$,

$$\int_0^\infty \cos cy\, dy \int_0^\infty e^{-ax} x^{n-1} \cos(xy)\, dx = \frac{\pi}{2} e^{-ac} c^{n-1}, \quad c \geqq 0.$$

Let $F(x, y) = e^{-ax} x^{n-1} \cos(xy)\cos(cy)$, and consider the integral

$$\int_0^\lambda dy \int_0^\infty F(x, y)dx, \quad \lambda > 0.$$

By equation (6), p. 471, of the *Elementary Treatise*,

$$\int_0^\infty F(x, y)dx = \Gamma(n) \cos n\theta \cos cy\, (y^2+a^2)^{-\frac{n}{2}},$$

where $\tan\theta = y/a$, $-\frac{1}{2}\pi < \theta < \frac{1}{2}\pi$; the integral converges uniformly for the range $0 \leqq y \leqq \lambda$, as may be readily verified. Next

$$\int_0^\lambda F(x, y)dy = \tfrac{1}{2} e^{-ax} x^{n-1} \left\{ \frac{\sin\lambda(x-c)}{x-c} + \frac{\sin\lambda(x+c)}{x+c} \right\},$$

and this integral converges uniformly for the unlimited range $x \geqq 0$, since $a>0$ and $n \geqq 1$.

Hence, by Theorem I of § 160, the order of integration may be changed, and therefore

$$\int_0^\lambda \cos(cy)\, dy \int_0^\infty e^{-ax} x^{n-1} \cos(xy)\, dx = \int_0^\infty e^{-ax} x^{n-1} dx \int_0^\lambda \cos(cy)\cos(xy)\, dy,$$

that is $$= \tfrac{1}{2}\int_0^\infty e^{-ax} x^{n-1} \frac{\sin\lambda(x-c)}{x-c}\, dx + \tfrac{1}{2}\int_0^\infty e^{-ax} x^{n-1} \frac{\sin\lambda(x+c)}{x+c}\, dx.$$

Now if $e^{-ax} x^{n-1} = f(x)$, the conditions required by the theorem of Ex. 4, § 159, are satisfied, since $f(x)$ has at most only one turning value (given by $x=(n-1)/a$). Hence, when $\lambda \to \infty$, the last two integrals tend to $\frac{\pi}{2} f(c)$ both for $c>0$ and for $c=0$. Hence

$$\int_0^\infty \cos(cy)\, dy \int_0^\infty e^{-ax} x^{n-1} \cos(xy)\, dx = \frac{\pi}{2} e^{-ac} c^{n-1}, \quad c \geqq 0.$$

Ex. 2. Prove that, if $a>0$ and $n \geqq 1$,

$$\int_0^\infty \sin(cy)\, dy \int_0^\infty e^{-ax} x^{n-1} \sin(xy)\, dx = \frac{\pi}{2} e^{-ac} c^{n-1}, \quad c > 0.$$

The proof is practically the same as in the case of Ex. 1, and may be left as an exercise.

These two examples are particular cases of Fourier's Double Integral. See *E.T.* § 194, pp. 499-501. It may be noted that the statement in that article (p. 501), that the absolute convergence of the integral $\int \frac{f(x)}{x}\, dx$ is sufficient for the validity of the transformation on p. 501, is not correct. See the articles by Pringsheim, *Math. Ann.* vol. 68 (year 1910) and vol. 71 (year 1912).

Ex. 3. The variables $x_1, x_2, \ldots, x_n$ are changed to $y_1, y_2, \ldots, y_n$ by the equations

$$x_1+x_2+\ldots+x_n=y_1,\ x_1+x_2+\ldots+x_{n-1}=y_1\,y_2, \ldots$$
$$x_1+x_2+\ldots+x_{n-r}=y_1\,y_2\ldots y_{r+1}, \ldots x_1=y_1\,y_2\ldots y_n.$$

If the variables $x_1, x_2, \ldots x_n$ may each take every real value that is not negative show that $y_2, y_3, \ldots, y_n$ will take every value between 0 and 1 (0 and 1 included), while y_1 takes every value from 0 to ∞. Further, show that the integral

$$\int_0^\infty \int_0^\infty \ldots \int_0^\infty F(x_1, x_2, \ldots x_n)dx_1\,dx_2\ldots dx_n$$

becomes

$$\int_0^\infty dy_1 \int_0^1 dy_2 \ldots \int_0^1 dy_n\ F_1(y_1, y_2, \ldots y_n)y_1^{n-1}y_2^{n-2}\ldots y_{n-1},$$

where $$F_1(y_1, y_2, \ldots y_n)=F(x_1, x_2, \ldots x_n).$$

We have

$$x_n=y_1(1-y_2),\ x_{n-1}=y_1\,y_2(1-y_3)\ldots,\ x_{n-r}=y_1\,y_2\ldots y_{r+1}(1-y_{r+2}), \ldots$$

so that $y_2, y_3, \ldots, y_n$ lie between 0 and 1 (including 0 and 1). On the other hand y_1 may vary from 0 to ∞, since $x_1=y_1\,y_2\ldots y_n$ and x_1 varies from 0 to ∞.

From the values just found for $x_n, x_{n-1}, \ldots$ it is easy to see that

$$J=\frac{\partial(x_1, x_2, \ldots x_n)}{\partial(y_1, y_2, \ldots y_n)}=\pm y_1^{n-1}y_2^{n-2}\quad .\, y_{n-1},$$

and therefore $|J|=y_1^{n-1}y_2^{n-2}\ldots y_{n-1}$. The required transformation of the integral follows at once.

EXERCISES XX.

1. Prove that, if $c>0$,

$$\int_0^\infty e^{-cy}\,dy\int_0^\infty \frac{\sin(xy)dx}{e^{2\pi x}-1}=\int_0^\infty \frac{x\,dx}{(x^2+c^2)(e^{2\pi x}-1)}.$$

2. Prove that, if $a>0$,

$$\int_a^\infty dc\int_0^\infty \frac{x\,dx}{(x^2+c^2)(e^{2\pi x}-1)}=\int_0^\infty \frac{\tan^{-1}\left(\dfrac{x}{a}\right)dx}{e^{2\pi x}-1}$$

3. Prove that, if $c>0$,

$$\int_0^\infty \frac{\cos x-e^{-cx}}{x}dx=\int_0^\infty (\cos x-e^{-cx})dx\int_0^\infty e^{-xy}\,dy=\log c.$$

4. If m and n are positive integers and $n \geqq m$, prove that

$$\int_0^\infty y^{m-1}\,dy\int_0^\infty e^{-xy}\sin^n x\,dx=(m-1)!\int_0^\infty \frac{\sin^n x}{x^m}dx.$$

5. If $u_n=\int_0^\infty e^{-xy}\sin^n x\,dx$, where n is a positive integer and $y>0$, prove that
$$(n^2+y^2)u_n=n(n-1)u_{n-2},$$
and then show that
$$u_n=\frac{n!}{y(2^2+y^2)(4^2+y^2)\dots(n^2+y^2)},\ n \text{ even},$$
$$u_n=\frac{n!}{(1^2+y^2)(3^2+y^2)\dots(n^2+y^2)},\ n \text{ odd}.$$

6. Deduce from Ex. 4 and Ex. 5, m and n being positive integers and $n\geqq m$, that
$$\int_0^\infty \frac{\sin^n x}{x^m}\,dx=\frac{1}{(m-1)!}\int_0^\infty y^{m-1}u_n\,dy.$$
If $m=2$, $n=3$, the integral is equal to $\frac{3}{4}\log 3$. Verify independently for these and other small numbers. (Bertrand.)

7. If $a>0$, $b>0$, $m\geqq 1$, $n\geqq 1$, $\theta=\tan^{-1}(y/a)$, $\varphi=\tan^{-1}(y/b)$, prove, by applying equation (6) (*E.T.* p. 471) and § 163, Ex. 1, that

(i) $$\int_0^\infty \frac{\cos m\theta\cos(xy)}{(a^2+y^2)^{\frac{m}{2}}}\,dy=\frac{\pi}{2\Gamma(m)}e^{-ax}x^{m-1},\ x>0;$$

(ii) $$\int_0^\infty \frac{\cos m\theta\cos n\varphi}{(a^2+y^2)^{\frac{m}{2}}(b^2+y^2)^{\frac{n}{2}}}\,dy=\frac{\pi}{2(a+b)^{m+n-1}}\,\frac{\Gamma(m+n-1)}{\Gamma(m)\Gamma(n)}.$$

8. Show that equations (i) and (ii) of Ex. 7 hold when the sine is substituted in place of the cosine, and deduce that
$$\int_0^\infty \frac{\cos(m\theta-n\varphi)\,dy}{(a^2+y^2)^{\frac{m}{2}}(b^2+y^2)^{\frac{n}{2}}}=\frac{\pi}{(a+b)^{m+n-1}}\,\frac{\Gamma(m+n-1)}{\Gamma(m)\Gamma(n)}.$$

9. By taking $a=b=1$ in Ex. 7 and Ex. 8, show that

(i) $$\int_0^{\frac{\pi}{2}} \cos m\theta\cos n\theta\cos^{m+n-2}\theta\,d\theta=\frac{\pi}{2^{m+n}}\,\frac{\Gamma(m+n-1)}{\Gamma(m)\Gamma(n)};$$

(ii) $$\int_0^{\frac{\pi}{2}} \sin m\theta\sin n\theta\cos^{m+n-2}\theta\,d\theta=\frac{\pi}{2^{m+n}}\,\frac{\Gamma(m+n-1)}{\Gamma(m)\Gamma(n)};$$

(iii) $$\int_0^{\frac{\pi}{2}} \cos(m-n)\theta\cos^{m+n-2}\theta\,d\theta=\frac{\pi}{2^{m+n-1}}\,\frac{\Gamma(m+n-1)}{\Gamma(m)\Gamma(n)}.$$

10. Show that, if $m>1$ and $0<n<1$,
$$\int_0^\infty y^{n-1}\,dy\int_0^\infty e^{-x}x^{m-1}\cos(xy)\,dx=\int_0^\infty e^{-x}x^{m-1}\,dx\int_0^\infty y^{n-1}\cos(xy)\,dy.$$
Deduce that

(i) $$\int_0^{\frac{\pi}{2}} \cos m\theta\cos^{m-n-1}\theta\sin^{n-1}\theta\,d\theta=\frac{\Gamma(n)\Gamma(m-n)}{\Gamma(m)}\cos\frac{n\pi}{2};$$

(ii) $$\int_0^{\frac{\pi}{2}} \sin m\theta\cos^{m-n-1}\theta\sin^{n-1}\theta\,d\theta=\frac{\Gamma(n)\Gamma(m-n)}{\Gamma(m)}\sin\frac{n\pi}{2}.$$

11. If a, b, α, β are all positive, then

$$\Gamma(a)\int_0^\infty \frac{e^{-\beta x}x^{b-1}\,dx}{(x+\alpha)^a}=\Gamma(b)\int_0^\infty \frac{e^{-\alpha x}x^{a-1}\,dx}{(x+\beta)^b}.$$

12. If $0 \leqq \alpha \leqq \frac{\pi}{2}$, show that

$$\int_0^\infty\int_0^\infty \frac{dx\,dy}{(x^2+2xy\cos\alpha+y^2+a^2)^2}=\frac{\alpha}{2a^2\sin\alpha},$$

and that, if $w=x^2+2xy\cos\alpha+y^2$,

(i) $\displaystyle\int_0^\infty\int_0^\infty e^{-w}\,dx\,dy=\frac{\alpha}{2\sin\alpha}$; (ii) $\displaystyle\int_0^\infty\int_0^\infty xy\,e^{-w}\,dx\,dy=\frac{\sin\alpha-\alpha\cos\alpha}{4\sin^3\alpha}$.

13. If

$$U=5x^2-2xy+2y^2+2x+2y+1,$$
$$V=6x^2+3y^2+4x+4y+2,$$

and if $\xi=x+y+1,\ \eta=y-2x,$

show that $U=\xi^2+\eta^2,\ V=2\xi^2+\eta^2,$

and then prove that $\displaystyle\int_{-\infty}^\infty\int_{-\infty}^\infty Ve^{-U}dx\,dy=\frac{\pi}{2}$.

[Change the variables to ξ and η. For the general transformation when U and V are real quadratic forms (positive and definite) see Hilton's *Linear Substitutions*, p. 75.]

14. $$\int_{-\infty}^\infty\int_{-\infty}^\infty \frac{dx\,dy}{(x^2+y^2+a^2)^{\frac{3}{2}}(x^2+y^2+b^2)^{\frac{1}{2}}}=\frac{2\pi}{a(a+b)},\quad a>0,\ b>0.$$

15. If $lm'+l'm$ is not zero, show that, the integral being assumed to converge,

$$\int_{-\infty}^\infty\int_{-\infty}^\infty F(lx+my,\,l'x-m'y)\,dx\,dy=|(lm'+l'm)^{-1}|\int_{-\infty}^\infty\int_{-\infty}^\infty F(x,y)\,dx\,dy.$$

16. If $a>0$ and $b>0$, show that

$$\int_0^\infty\int_0^\infty \varphi(a^2x^2+b^2y^2)\,dx\,dy=\frac{\pi}{4ab}\int_0^\infty \varphi(x)\,dx,$$

the integral being assumed to be convergent.

17. If $a>0$, $b>0$, $m>0$, $n>0$, show that

$$\int_0^\infty\int_0^\infty e^{-(ax^2+by^2)}x^{2m-1}y^{2n-1}\,dx\,dy=\frac{\Gamma(m)\Gamma(n)}{4a^m b^n},$$

and then deduce that

$$\int_0^{\frac{\pi}{2}} \frac{(\cos\theta)^{2m-1}(\sin\theta)^{2n-1}\,d\theta}{(a\cos^2\theta+b\sin^2\theta)^{m+n}}=\frac{1}{2a^m b^n}\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

CHAPTER XV

INTEGRATION OF SERIES. GAMMA FUNCTIONS

164. Integration of Series. When the terms of a series $\Sigma u_n(x)$ are not bounded, or when the range of integration is not finite, the integration of the series falls within the region of the improper integral. The methods of dealing with the improper integral that have been explained in previous chapters are applicable also to such series, but in many cases of practical importance it is possible to integrate the series by applying elementary theorems, as in the following examples.

Ex. 1. Show that $\int_0^1 \frac{\log x}{1-x}dx = -\frac{\pi^2}{6}$.

Here we may write

$$\frac{1}{1-x}=\sum_{n=0}^{N-1} x^n + \frac{x^N}{1-x},$$

and therefore
$$\int_0^1 \frac{\log x\, dx}{1-x} = -\sum_{n=1}^{N}\frac{1}{n^2}+\int_0^1 \frac{x^N \log x}{1-x}dx.$$

Now $x \log x/(1-x)$ is *bounded*, for, if $x=e^{-y}$,

$$\frac{x\log x}{1-x}=\frac{-y\,e^{-y}}{1-e^{-y}}=\frac{-y}{e^y-1}=\frac{-1}{1+\frac{1}{2}y+\frac{1}{6}y^2+\dots},$$

so that
$|x\log x/(1-x)| \leqq 1$ if $0 \leqq x \leqq 1$. Hence

$$\left|\int_0^1 \frac{x^N\log x}{1-x}dx\right| \leqq \int_0^1 x^{N-1}dx=\frac{1}{N},$$

and therefore
$$\int_0^1 \frac{\log x\, dx}{1-x}=-\sum_{n=1}^{\infty}\frac{1}{n^2}=-\frac{\pi^2}{6}.$$

Ex. 2.
$$\int_0^{\infty}\frac{\sin cx\, dx}{e^x-1}=\frac{1}{2}\left(\pi \coth \pi c-\frac{1}{c}\right),$$

if c is real and not zero, or, if c is complex, $c=a+ib$, $|b|<1$.

If $x>0$, we have

$$\frac{1}{e^x-1}=\frac{e^{-x}}{1-e^{-x}}=\sum_{n=1}^{N}e^{-nx}+\frac{e^{-Nx}}{e^x-1}. \quad \dots\dots\dots\dots(1)$$

Now the function $\sin cx/(e^x - 1)$ is bounded for the unlimited range $x \geqq 0$ when c satisfies the conditions stated.

(i) $\sin cx/(e^x - 1) \to c$ when $x \to 0$, and therefore, when x is small, say when $0 \leqq |x| \leqq k$, $|\sin cx/(e^x - 1)| < c'$, where c' differs little from $|c|$.

(ii) If $x > k$ and c real, $|\sin cx/(e^x - 1)| < 1/(e^k - 1)$, while if c is complex, $c = a + ib$, $|\sin cx| \leqq \cosh bx + \sinh bx \leqq e^{b'x}$, where $b' = |b|$, and therefore, if $b' < 1$,

$$e^x - 1 \quad < \frac{e^{b'x}}{e^x - 1} < \frac{1}{e^{(1-b')x} - 1} < \frac{1}{e^{(1-b')k} - 1}.$$

There is thus a constant K such that $|\sin cx/(e^x - 1)| < K$ if $x \geqq 0$, and therefore, multiplying equation (1) by $\sin cx$ and integrating, we find

$$\int_0^\infty \frac{\sin cx\, dx}{e^x - 1} = \sum_{n=1}^{N} \frac{c}{c^2 + n^2} + R_N,$$

where
$$|R_N| < K \int_0^\infty e^{-Nx}\, dx = \frac{K}{N},$$

and therefore $R_N \to 0$ when $N \to \infty$. Hence (§ 94, (1)),

$$\int_0^\infty \frac{\sin cx\, dx}{e^x - 1} = \sum_{n=1}^{\infty} \frac{c}{c^2 + n^2} = \tfrac{1}{2}\left(\pi \coth \pi c - \frac{1}{c}\right).$$

Cor. 1. Let $x = 2\pi y$, $2\pi c = \alpha$, then

$$\int_0^\infty \frac{\sin \alpha y\, dy}{e^{2\pi y} - 1} = \frac{1}{2}\left(\frac{1}{e^\alpha - 1} - \frac{1}{\alpha} + \frac{1}{2}\right).$$

where $\alpha \gtrless 0$ if α is real, or, when α is complex ($\alpha = \beta + i\gamma$), $|\gamma| < 2\pi$.

Cor. 2. Since $\dfrac{1}{e^x + 1} = \dfrac{1}{e^x - 1} - \dfrac{2}{e^{2x} - 1}$,

$$\int_0^\infty \frac{\sin cx\, dx}{e^x + 1} = \frac{1}{2}\left(\frac{1}{c} - \frac{\quad}{\sinh \pi c}\right).$$

Cor. 3. The above formulae are valid for $c = 0$ in the sense that the limits of the integral and of its value tend to 0 when $c \to 0$.

165. General Theorems. In the following theorems it is assumed that each term $u_n(x)$ of a series $\Sigma u_n(x)$ is integrable over the range $a \leqq x \leqq b$ or over the unlimited range $x \geqq a$ according as the range of integration is finite or infinite. Further, $u_n(x)$ will be taken as a product, $f(x)v_n(x)$, where $f(x)$ is independent of n and $f(x)$ and $v_n(x)$ are integrable.

THEOREM I. *If* (i) *the series* $\Sigma v_n(x)$ *converges uniformly for the range* $a \leqq x \leqq b$, *and* (ii) *the integral* $\int_a^b |f(x)|\,dx$ *converges, then*

$$\int_a^b \left[\sum_0^\infty f(x) v_n(x)\right] dx = \sum_0^\infty \left[\int_a^b f(x) v_n(x)\, dx\right].$$

The upper limit b may be finite or infinite; if $b = \infty$ the series $\Sigma v_n(x)$ is to converge uniformly for the unlimited range $x \geqq a$.

Let $$R_m(x)=\sum_{n=m+1}^{\infty} v_n(x),\quad A=\int_a^b |f(x)|dx.$$

By condition (i) N may be chosen so that, if $a \leqq x \leqq b$ and $m \geqq N$, we shall have $|R_m(x)|<\varepsilon/A$, where ε is arbitrarily small, and therefore

$$\left|\int_a^b f(x)\,R_m(x)\,dx\right|<\varepsilon,\quad \underset{m\to\infty}{\mathcal{L}} \int_a^b f(x)\,R_m(x)\,dx=0,$$

so that the integral of $\Sigma f(x)\,v_n(x)$ over $(a,\,b)$ converges.

Next,

$$\sum_0^m\left[\int_a^b f(x)\,v_n(x)\,dx\right]=\int_a^b\left[\sum_0^m f(x)\,v_n(x)\right]dx$$
$$=\int_a^b\left[\sum_0^{\infty} f(x)\,v_n(x)\right]dx-\int_a^b f(x)\,R_m(x)\,dx.$$

Let $m\to\infty$ and the result follows. The proof is the same if $b=\infty$.

Ex. 1. $\displaystyle\int_0^1 \log x \log(1+x)dx=2-2\log 2-\frac{1}{12}\pi^2.$

The series for $\log(1+x)$ converges uniformly for the range $0\leqq x\leqq 1$ and $\displaystyle\int_0^1 |\log x|\,dx=1$. Hence the integral is equal to

$$\sum_1^{\infty}\frac{(-1)^n}{n(n+1)^2}=\sum_1^{\infty}(-1)^n\left[\frac{1}{n}-\frac{1}{n+1}-\frac{1}{(n+1)^2}\right]=2-2\log 2-\frac{1}{12}\pi^2.$$

THEOREM II. *If $f(x)$ and $v_n(x)\,(n=0,\,1,\,2\,\ldots)$ are positive (or zero) for the range $a\leqq x\leqq b$, and if the series $\Sigma v_n(x)$ is only uniformly convergent in general for that range, then the integral*

$$\int_a^b\left[\sum_0^{\infty} f(x)\,v_n(x)\right]dx \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(\alpha)$$

and the series

$$\sum_0^{\infty}\left[\int_a^b f(x)\,v_n(x)\,dx\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(\beta)$$

are equal, provided that either *the integral* (α) or *the series* (β) *is determinate.*

Case 1. The series $\Sigma v_n(x)$ ceases to converge uniformly at b and only at b; it may or may not converge when $x=b$.

Suppose first that the integral (α) is determinate and equal, say, to A; it is therefore possible to choose $\lambda\ (<b)$ so that we shall have

$$R_\lambda=\int_\lambda^b\left[\sum_0^{\infty} f(x)\,v_n(x)\right]dx<\varepsilon,\quad \underset{\lambda\to b}{\mathcal{L}} R_\lambda=0,$$

since $f(x)$ and $v_n(x)$ are positive and ε is an arbitrarily small

positive number. Now, if $a<\lambda<b$, Theorem I is applicable; hence

$$\sum_0^\infty\left[\int_a^\lambda f(x)v_n(x)\,dx\right]=\int_a^\lambda\left[\sum_0^\infty f(x)v_n(x)\right]dx$$
$$=A-R_\lambda.$$

Let $\lambda\to b$; then $R_\lambda\to 0$ while the series on the left of the equation tends to the series (β), so that the series (β) is equal to the integral (α).

Next, suppose that the series (β) is determinate and equal, say, to B. By Theorem I, if $a<\lambda<b$,

$$\int_a^\lambda\left[\sum_0^\infty f(x)v_n(x)\right]dx=\sum_0^\infty\left[\int_a^\lambda f(x)v_n(x)\,dx\right]<B,$$

since $f(x)$ and $v_n(x)$ are positive (when not zero); thus the integral

$$\int_a^\lambda\left[\sum_0^\infty f(x)v_n(x)\right]dx$$

is a positive, monotonic, increasing function of λ which is less than B and therefore, when $\lambda\to b$, tends to a limit which is not greater than B. In other words the integral (α) converges.

Again, since the series (β) converges,

$$\mathop{\mathcal{L}}_{m\to\infty}\sum_{n=m+1}^\infty\left[\int_a^b f(x)v_n(x)\,dx\right]=0.$$

Hence, since the functions $f(x)v_n(x)$ are integrable,

$$\int_a^b\left[\sum_0^m f(x)v_n(x)\right]dx=\sum_0^m\left[\int_a^b f(x)v_n(x)\,dx\right]$$
$$=B-\sum_{n=m+1}^\infty\left[\int_a^b f(x)v_n(x)\,dx\right].$$

Now let $m\to\infty$, and it follows that the integral (α) is equal to B; that is, the integral (α) and the series (β) are equal.

Case 2. If the series $\Sigma v_n(x)$ ceases to converge uniformly at a and only at a, a very slight modification of the above proof shows that the theorem is true in this case; it follows then in the usual way that the theorem is true when $\Sigma v_n(x)$ is only uniformly convergent in general.

Cor. For brevity, omit the letter x in the functional symbols

$f(x)$, $v_n(x)$. Then, if $f(x)$ and $v_n(x)$ are not always positive, the product $f(x)v_n(x)$ may be expressed as the sum

$$(f+|f|)\cdot(v_n+|v_n|)-|f|\cdot(v_n+|v_n|)-(f+|f|)\cdot|v_n|+|f|\cdot|v_n|,$$

and each of the four functions in this sum is of the form $g(x)w_n(x)$, where $g(x)$ and $w_n(x)$ are positive (or zero). Hence Theorem II is true if either the integral

$$\int_a^b\left[\sum_0^\infty |f(x)|\cdot|v_n(x)|\right]dx$$

or the series
$$\sum_0^\infty\left[\int_a^b |f(x)|\cdot|v_n(x)|dx\right]$$

is determinate.

THEOREM III. *If $f(x)$ and $v_n(x)$ $(n=0, 1, 2, \ldots)$ are positive (or zero) for the unlimited range $x \geqq a$, and if the series $\Sigma v_n(x)$ is uniformly convergent in general for the arbitrarily large range $a \leqq x \leqq b$, then the integral*

$$\int_a^\infty\left[\sum_0^\infty f(x)v_n(x)\right]dx \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots(\alpha)$$

and the series
$$\sum_0^\infty\left[\int_a^\infty f(x)v_n(x)dx\right] \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots(\beta)$$

are equal, provided that either *the integral* (α) or *the series* (β) *is determinate.*

The proof of this theorem when $\Sigma v_n(x)$ converges uniformly in the arbitrarily large interval (a, b) follows so closely the lines of the proof of Case 1 of Theorem II that its detailed statement may be left to the student. The modifications required when the series only converges uniformly in general in (a, b) have been dealt with in the proof of Theorem II. Thus, if a is the only point of non-uniform convergence of the series $\Sigma v_n(x)$, take $c>a$; then Theorem II applies when the interval of integration is (a, c), and the proof of Theorem III applies when the convergence of the series $\Sigma v_n(x)$ is uniform in (c, b), so that Theorem III holds when a is the only point of non-uniform convergence of $\Sigma v_n(x)$.

Cor. The Corollary of Theorem II is also true for Theorem III; in the integrals of the Corollary of Theorem II we merely put $b=\infty$.

Theorems II and III are due to Hardy, *Messenger of Mathematics*, Vol. 35 (year 1905), pp. 126-130. See also Bromwich, *Infinite Series* (2nd Ed.), pp. 495-502.

Examples 1 and 2 of § 164 illustrate these theorems. Thus in Ex. 1 the series for $(1-x)^{-1}$ consists of positive terms and converges uniformly for $0 \leqq x \leqq \lambda < 1$, and the series (if $\log(1/x)$ instead of $\log x$ is taken) converges. Similarly in Example 2 the series Σe^{-nx} converges uniformly if $x \geqq \lambda > 0$, and the Corollary of Theorem III applies.

Another theorem due to Dini (*Fondamenti*, p. 391) may be given. If the series is $\Sigma u_n(x)$ and if $w_n(\xi)$ denote the integral

$$\int_a^\xi u_n(x)\,dx$$

the theorem may be stated as follows:

THEOREM IV. *If* (i) *the series $\Sigma u_n(x)$ converges uniformly for the arbitrarily large range $a \leqq x \leqq b$, and* (ii) *the series $\Sigma w_n(\xi)$ converges uniformly for the unlimited range $\xi \geqq a$, then the integral*

$$\int_a^\infty \left[\sum_0^\infty u_n(x)\right] dx \quad \ldots\ldots\ldots\ldots(\alpha)$$

and the series

$$\sum_0^\infty \left[\int_a^\infty u_n(x)\,dx\right] \quad \ldots\ldots\ldots\ldots(\beta)$$

are each determinate, and the integral (α) *is equal to the series* (β).

By condition (i)

$$\int_a^\xi \left[\sum_0^\infty u_n(x)\right] dx = \sum_0^\infty w_n(\xi).$$

Again, by condition (ii), $\Sigma w_n(\xi)$ is a continuous function of ξ for $\xi \geqq a$, and therefore

$$\mathop{\mathcal{L}}_{\xi\to\infty} \left[\sum_0^\infty w_n(\xi)\right] = \sum_0^\infty \left[\mathop{\mathcal{L}}_{\xi\to\infty} w_n(\xi)\right];$$

that is

$$\int_a^\infty \left[\sum_0^\infty u_n(x)\right] dx = \sum_0^\infty \left[\int_a^\infty u_n(x)\,dx\right].$$

The following additional examples illustrate the theorems.

Ex. 2. $\displaystyle\int_0^1 \log\left(\frac{1+x}{1-x}\right)\frac{dx}{x} = \frac{\pi^2}{4}.$

Here

$$\log\{(1+x)/(1-x)\} = 2\sum_1^\infty \frac{x^{2n-1}}{2n-1},$$

and the series converges uniformly for the range $0 \leqq x \leqq \lambda < 1$. By Th. II, since

$$\int_0^1 \left[2\sum_1^\infty \frac{x^{2n-1}}{2n-1}\right]\frac{dx}{x} = 2\sum_1^\infty \frac{1}{(2n-1)^2} = \frac{\pi^2}{4},$$

the result follows.

Ex. 3. $$\int_0^\infty\left[\sum_{n=1}^\infty \frac{xe^{-\alpha x}}{4n^2\pi^2+x^2}\right]dx=\int_0^\infty \frac{x\,dx}{(\alpha^2+x^2)(e^{2\pi x}-1)},$$
if α is real and positive, or, if $\alpha=\beta+i\gamma$ and $\beta>0$.

Here $v_n(x)=(4n^2\pi^2+x^2)^{-1}$ and $\Sigma v_n(x)$ converges uniformly for the unlimited range $x\geqq 0$, since $\Sigma v_n(x)\leqq \Sigma(4n^2\pi^2)^{-1}=1/24$. The given integral is therefore (by Theorem I) equal to

$$\sum_{n=1}^\infty\left[\int_0^\infty \frac{xe^{-\alpha x}\,dx}{4n^2\pi^2+x^2}\right]=\sum_1^\infty\left[\int_0^\infty \frac{ye^{-2n\pi\alpha y}\,dy}{1+y^2}\right],$$

by putting $x=2n\pi y$.

Now the series $\Sigma|e^{-2n\pi\alpha y}|$ or $\Sigma e^{-2n\pi\beta y}$, when α is complex, converges uniformly in general in an arbitrarily large interval $(0, b)$, and by Theorem III or Theorem III Cor., the summation and integration may be interchanged, if the integral thus obtained is convergent. Hence the given integral is equal to

$$\int_0^\infty\left[\sum_1^\infty e^{-2n\pi\alpha y}\right]\frac{y\,dy}{1+y^2}=\int_0^\infty \frac{y\,dy}{(1+y^2)(e^{2\pi\alpha y}-1)},$$

and if now the variable of integration is changed to x, where $x=\alpha y$, the integral is that stated above, since the integral converges.

Note that $$\sum_{n=1}^\infty \frac{1}{4n^2\pi^2+x^2}=\tfrac{1}{2}\left(\frac{1}{e^x-1}-\frac{1}{x}+\frac{1}{2}\right)e^{-\alpha x},$$

by § 94, when $\tfrac{1}{2}x$ is put in place of x in equation (1).

Ex. 4. $$\int_0^\infty\left[\sum_1^\infty \frac{e^{-\alpha x}}{4n^2\pi^2+x^2}\right]dx=\int_0^\infty \frac{\tan^{-1}(x/\alpha)\,dx}{e^{2\pi x}-1}$$

where α is the same as in Ex. 3.

As in Ex. 3 the given integral is equal to

$$\sum_1^\infty\left[\int_0^\infty \frac{e^{-\alpha x}\,dx}{4n^2\pi^2+x^2}\right]=\frac{1}{2\pi}\sum_1^\infty\left[\frac{1}{n}\int_0^\infty \frac{e^{-2n\pi\alpha y}\,dy}{1+y^2}\right].$$

Again, as in Ex. 3, the order of summation and integration may be interchanged, providing the integral thus obtained converges. The change of order gives

$$\frac{1}{2\pi}\int_0^\infty\left[\sum_1^\infty \frac{1}{n}e^{-2n\pi\alpha y}\right]\frac{dy}{1+y^2}=\frac{1}{2\pi}\int_0^\infty -\log(1-e^{-2\pi\alpha y})\frac{dy}{1+y^2},$$

and, after integration by parts and the substitution of x for αy, the integral becomes
$$2\pi\int_0^\infty \frac{\tan^{-1}(x/\alpha)\,dx}{e^{2\pi x}-1},$$

where $\tan^{-1}(x/\alpha)=0$ when $x=0$, the integral being convergent. Hence the given equation.

From the value of the series stated at the end of Ex. 3 we have

$$\int_0^\infty\left(\frac{1}{e^x-1}-\frac{1}{x}+\frac{1}{2}\right)\frac{e^{-\alpha x}}{x}\,dx=2\int_0^\infty \frac{\tan^{-1}(x/\alpha)}{e^{2\pi x}-1}\,dx.$$

Ex. 5. If B_n is Bernoulli's Number, show that

$$\frac{B_n}{4n}=\int_0^\infty \frac{y^{2n-1}\,dy}{e^{2\pi y}-1}.$$

Let $$f(t)=\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2},\ t \text{ real};$$

then, by § 94, (5),

$$f(t)=\sum_{n=1}(-1)^{n-1}\frac{B_n}{(2n)!}t^{2n-1}, \quad\text{.......................(1)}$$

and, by § 164, Ex. 2,

$$f(t)=2\int_0^\infty \frac{\sin ty\,dy}{e^{2\pi y}-1}. \quad\text{..................................(2)}$$

Now $$\frac{\sin ty}{e^{2\pi y}-1}=\frac{y}{e^{2\pi y}-1}\cdot\frac{\sin ty}{y},$$

and the series $$\sum_1^\infty(-1)^{n-1}\frac{t^{2n-1}y^{2n-2}}{(2n-1)!}$$

converges uniformly with respect to y in an arbitrarily large interval $(0, b)$. Hence, expressing $\sin ty$ by a power series in (ty), and interchanging the order of summation and integration, we find

$$f(t)=2\sum_1^\infty(-1)^{n-1}\frac{t^{2n-1}}{(2n-1)!}\int_0^\infty\frac{y^{2n-1}\,dy}{e^{2\pi y}-1},$$

and therefore, by equating the coefficients of t^{2n-1} in (1) and (2), the stated value of B_n is obtained.

EXERCISES XXI.

Many of the series required in this set will be found in or may be derived from the Examples and Exercises in Chapter VIII.

1. $\displaystyle\int_0^1\frac{(\log x)^{2r-1}}{1-x}dx=-(2r-1)!\left[\sum_{n=1}^\infty\frac{1}{n^{2r}}\right]=-\frac{2^{2r-1}}{2r}B_r\pi^{2r}.$*

2. $\displaystyle\int_0^1\frac{(\log x)^{2r-1}}{1+x}dx=-(2r-1)!\left[\sum_{n=1}^\infty\frac{(-1)^{n-1}}{n^{2r}}\right]=-\frac{2^{2r-1}-1}{2r}B_r\pi^{2r}.$

3. $\displaystyle\int_0^1\frac{x^p\log x\,dx}{(1+x)^2}=-\left[\frac{1}{(p+1)^2}-\frac{2}{(p+2)^2}+\frac{3}{(p+3)^2}-\dots\right],\ p+1>0.$

4. $\displaystyle\int_0^\infty\frac{x\log x\,dx}{(1+x^2)^3}=\frac{1}{8}\int_0^1 x\log\left(\frac{1-x}{1+x}\right)dx=-\frac{1}{8}.$

5. If $|a|<1$ and n a positive integer,

(i) $\displaystyle\int_0^{2\pi}\frac{\cos nx\,dx}{1-2a\cos x+a^2}=\frac{2\pi a^n}{1-a^2};$

(ii) $\displaystyle\int_0^{2\pi}\frac{(1+2\cos x)^n\cos nx\,dx}{1-2a\cos x+a^2}=\frac{2\pi}{1-a^2}(1+a+a^2)^n;$

(iii) $\displaystyle\int_0^{2\pi}\frac{(b+2c\cos x)^n\cos nx\,dx}{1-2a\cos x+a^2}=\frac{2\pi}{1-a^2}(c+ba+ca^2)^n.$

* For the value of B_r see § 94.

6. $\int_0^{2\pi}\frac{\sin^2 x\cos x\,dx}{(1-2a\cos x+a^2)^2}=\frac{\pi a}{1-a^2},\ a^2<1.$

7. $\int_0^{\frac{\pi}{2}}\log(a^2\sin^2 x+b^2\cos^2 x)dx=\pi\log\left(\frac{a+b}{2}\right),\ a>0,\ b>0,$

and show that the result holds if $a=0,\ b\neq 0$, or if $a\neq 0,\ b=0$.

Deduce that

$$\int_0^{\infty}\log\left(1+\frac{a^2}{x^2}\right)\frac{b\,dx}{b^2+x^2}=\pi\log\left(1+\frac{a}{b}\right)=2\int_0^{\infty}\frac{x\tan^{-1}\left(\frac{a}{x}\right)d[illegible]}{x^2+b^2}$$

8. $\int_0^{\frac{\pi}{2}}\frac{1-a\cos 2x}{1-2a\cos 2x+a^2}\log(\cos x)\,dx=\frac{\pi}{4}\log\frac{1+a}{4},\ a^2<1,$

$$=\frac{\pi}{4}\log\frac{a}{1+a},\ a^2>1.$$

9. (i) $\int_0^{\frac{\pi}{2}}(\log\tan x)^2\,dx=\int_0^{\frac{\pi}{2}}(\log\cot x)^2\,dx=\frac{\pi^3}{8};$

(ii) $\int_0^{\frac{\pi}{2}}(\log\sin x)^2\,dx=\int_0^{\frac{\pi}{2}}(\log\cos x)^2 dx=\frac{\pi}{2}\left\{(\log 2)^2+\frac{\pi^2}{12}\right\};$

(iii) $\int_0^{\frac{\pi}{2}}(\log\sin x)(\log\cos x)dx=\frac{\pi}{2}\left\{(\log 2)^2-\frac{\pi^2}{24}\right\}.$

10. $\int_0^1\log\left(\frac{1}{x}\right)\frac{dx}{2-x}=\frac{\pi^2}{12}-\frac{1}{2}(\log 2)^2.$

11. $\int_0^{\infty}\log\left(\frac{1+2a\sin x+a^2}{1-2a\sin x+a^2}\right)\frac{dx}{x}=2\pi\tan^{-1}a,\ a^2<1.$

12. $\int_0^{\infty}\log\left(\frac{1+2a\cos bx+a^2}{1+2a\cos cx+a^2}\right)\frac{dx}{x^2}=\frac{\pi a(c-b)}{1+a},\ a^2<1,\ b>0,\ c>0.$

13. $\int_0^{\infty}\left(\frac{\log x}{x-1}\right)^2dx=-\int_0^1(\log x)^3\frac{1+x}{(1-x)^3}dx=\pi^2.$*

14. (i) $\int_0^{\infty}\frac{x\,dx}{e^x-1}=\frac{\pi^2}{6};$ (ii) $\int_0^{\infty}\frac{x\,dx}{e^x+1}=\frac{\pi^2}{12};$

(iii) $\int_0^{\infty}\frac{x\,dx}{\sinh x}=\frac{\pi^2}{4}.$

15. (i) $\int_0^{\infty}\frac{\sin ax}{\sinh x}dx=\frac{\pi}{2}\tanh\frac{\pi a}{2};$ $\int_0^{\infty}\frac{x\cos ax}{\sinh x}$ [illegible] $\frac{\pi a}{2}.$

16. (i) $\int_0^{\infty}\frac{\cos ax}{\cosh x}dx=\frac{\pi}{2}\operatorname{sech}\frac{\pi a}{2};$ (ii) $\int_0^{\infty}\frac{\cosh ax}{\cosh x}=\frac{\pi}{2}\sec\frac{\pi a}{2},\ |a|<1.$

17. (i) $\int_0^{\infty}\frac{\sin ax}{\cosh x}\frac{dx}{x}=2\tan^{-1}\left(\tanh\frac{\pi a}{4}\right);$

(ii) $\int_0^{\infty}\frac{\sinh ax}{\cosh x}\frac{dx}{x}=\log\cot\frac{1-a}{4}\pi,\ |a|<1.$

* Examples 9-13 are taken from Wolstenholme's *Mathematical Problems*, where many examples of a similar type are to be found.

18. $\displaystyle\int_0^\infty \frac{\cosh ax}{\sinh \pi x}\sin cx\,dx = \tfrac12 \frac{\sinh c}{\cosh c + \cos a}$, $0 < a < \pi$.

19. If $f(x) = \dfrac{1}{e^x - 1} - \dfrac{1}{x} + \dfrac{1}{2}e^{-x}$, prove that

$$\int_0^\infty f(x)\frac{dx}{x} = \int_0^\infty f(2x)\frac{dx}{x},$$

and deduce that

$$\int_0^\infty \left(1 - \frac{x}{\sinh x}\right)\frac{dx}{x^2} = \log 2.$$

$\left[\text{Note that } [f(x) - f(2x)]\dfrac{1}{x} = \dfrac{1}{2}\dfrac{e^{-x} - e^{-2x}}{x} - \dfrac{1}{2}\left(1 - \dfrac{x}{\sinh x}\right)\dfrac{1}{x^2}.\right]$

20. If $a > 0$, $b > 0$, deduce from Ex. 19 that

$$\int_0^\infty \left(\frac{a}{\sinh ax} - \frac{b}{\sinh bx}\right)\frac{dx}{x} = (b - a)\log 2.$$

21. If $a^2 < 1$, $b^2 < 1$,

$$\int_0^\pi \frac{\sin^2 x\,dx}{(1 - 2a\cos x + a^2)(1 - 2b\cos x + b^2)} = \frac{\pi}{2(1 - ab)}.$$

22. If $-1 < p < 1$ and $-\pi < \lambda < \pi$,

$$\int_0^\infty \frac{x^{-p}\,dx}{1 + 2x\cos\lambda + x^2} = \int_0^1 \frac{(x^p + x^{-p})dx}{1 + 2x\cos\lambda + x^2} = \frac{\pi}{\sin p\pi}\,\frac{\sin p\lambda}{\sin\lambda};$$

and, if $0 < m < 2n$, show that

$$\int_0^\infty \frac{x^{m-1}\,dx}{x^{2n} + 2x^n\cos\lambda + 1} = \frac{\pi}{n\sin(m\pi/n)}\cdot\frac{\sin\{(n - m)\lambda/n\}}{\sin\lambda}.$$

23. If $a^2 < 1$ and $c > 0$, prove that

(i) $\displaystyle\int_0^\infty \frac{1}{1 + x^2}\,\frac{dx}{1 - 2a\cos cx + a^2} = \frac{\pi}{2}\cdot\frac{1}{1 - a^2}\cdot\frac{1 + ae^{-c}}{1 - ae^{-c}}$,

(ii) $\displaystyle\int_0^\infty \log(1 - 2a\cos cx + a^2)\frac{dx}{1 + x^2} = \pi\log(1 - ae^{-c})$;

(iii) $\displaystyle\int_0^\infty \frac{x\sin cx}{1 - 2a\cos cx + a^2}\,\frac{dx}{1 + x^2} = \frac{\pi}{2}\,\frac{1}{e^c - a}$.

24. (i) $\displaystyle\int_0^1\int_0^1 \frac{dx\,dy}{1 - xy} = \frac{\pi^2}{6}$;

(ii) $\displaystyle\int_0^1\int_0^1 \frac{dx\,dy}{(1 - k^2x^2y^2)\sqrt{(1 - x^2)}\sqrt{(1 - y^2)}}$

$$= \frac{\pi}{2}\int_0^1 \frac{dx}{\sqrt{(1 - x^2)}\sqrt{(1 - k^2x^2)}},\quad k^2 < 1.$$

25. If $\alpha > 0$, $\gamma - \alpha > 0$, $|x| < 1$, show that

$$F(\alpha, \beta, \gamma, x) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma - \alpha)}\int_0^1 t^{\alpha-1}(1 - t)^{\gamma-\alpha-1}\frac{dt}{(1 - xt)^\beta},$$

and deduce that, if also $\gamma - \alpha - \beta > 0$,

$$F(\alpha, \beta, \gamma, 1) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)}.$$

26. By changing the variable of integration in the integral in Ex. 25, prove that,

$$\begin{aligned}F(\alpha, \beta, \gamma, x) &= (1-x)^{-\alpha} F\left(\alpha, \gamma-\beta, \gamma, \frac{x}{x-1}\right)\\ &= (1-x)^{-\beta} F\left(\beta, \gamma-\alpha, \gamma, \frac{x}{x-1}\right)\\ &= (1-x)^{\gamma-\alpha-\beta}\, F(\gamma-\beta, \gamma-\alpha, \gamma, x)\\ &= (1-x)^{\gamma-\alpha-\beta}\, F(\gamma-\alpha, \gamma-\beta, \gamma, x).\end{aligned}$$

166. Integrals for Euler's Constant. The following Lemma is useful.

Lemma. If $f(t) = \dfrac{1}{e^t-1} - \dfrac{1}{t} + \dfrac{1}{2}$, then

(i) $0 < f(t) < \dfrac{1}{12}t$, when $t > 0$;

(ii) $f(t) \to 0$ and $\dfrac{1}{t}f(t) \to \dfrac{1}{12}$, when $t \to 0$.

By § 94, equation (2), if $t > 0$,

$$f(t) = 2t\sum_{n=1}^{\infty} \frac{1}{4n^2\pi^2 + t^2} < \frac{2t}{4\pi^2}\sum_1^{\infty}\frac{1}{n^2} = \frac{1}{12}t,$$

and the statements in (i) and (ii) follow at once.

Ex. 1. If $\gamma_n = 1 + \dfrac{1}{2} + \dfrac{1}{3} + \ldots + \dfrac{1}{n} - \log n$, prove that

(i) $\gamma_n = \displaystyle\int_0^{\infty}\left(\frac{1}{1-e^{-t}} - \frac{1}{t}\right)e^{-t}\,dt + \int_0^{\infty}\left(\frac{1}{t} - \frac{1}{e^t-1}\right)e^{-nt}\,dt$;

(ii) $\gamma = \underset{n\to\infty}{\mathcal{L}}\, \gamma_n = \displaystyle\int_0^{\infty}\left(\frac{1}{1-e^{-t}} - \frac{1}{t}\right)e^{-t}\,dt$,

where γ is Euler's Constant (§ 148, Ex. 7).

(i) $$\sum_{k=1}^{n}\frac{1}{k} = \sum_{k=1}^{n}\int_0^{\infty} e^{-kt}\,dt = \int_0^{\infty}\frac{1-e^{-nt}}{1-e^{-t}}\,e^{-t}\,dt.$$

Also $$\log n = \int_0^{\infty}\frac{e^{-t}-e^{-nt}}{t}\,dt.$$

These values give the equation (i).

(ii) By the Lemma

$$0 < \int_0^{\infty}\left(\frac{1}{e^t-1} - \frac{1}{t} + \frac{1}{2}\right)e^{-nt}\,dt < \frac{1}{12}\int_0^{\infty} te^{-nt}\,dt;$$

that is, $$0 < \int_0^{\infty}\left(\frac{1}{e^t-1} - \frac{1}{t}\right)e^{-nt}\,dt + \frac{1}{2n} < \frac{1}{12n^2},$$

so that $$\underset{n\to\infty}{\mathcal{L}}\int_0^{\infty}\left(\frac{1}{e^t-1} - \frac{1}{t}\right)e^{-nt}\,dt = 0,$$

from which the stated value of γ follows.

Ex. 2. Show that $\gamma=\int_0^\infty\left(\frac{1}{1+t}-e^{-t}\right)\frac{dt}{t}$.

Express the integrand in Ex. 1, (ii), in the form

$$\left\{\frac{1}{t(1+t)}-\frac{e^{-t}}{t}\right\}+\left\{\frac{e^{-t}}{1-e^{-t}}-\frac{1}{t(1+t)}\right\}.$$

Now $$\frac{e^{-t}}{1-e^{-t}}-\frac{1}{t(1+t)}=\frac{d}{dt}\log\left[\frac{(1-e^{-t})(1+t)}{t}\right],$$

and $\log[(1-e^{-t})(1+t)/t]$ tends to zero both when $t\to0$ and when $t\to\infty$, so that the integral for γ takes the form stated.

Ex. 3. Show that $\gamma=\int_0^\infty\left(\frac{1}{1+t^2}-e^{-t}\right)\frac{dt}{t}$.

$$\left(\frac{1}{1+t}-e^{-t}\right)\frac{1}{t}=\left(\frac{1}{1+t^2}-e^{-t}\right)\frac{1}{t}+\left(\frac{1}{1+t}-\frac{1}{1+t^2}\right)\frac{1}{t},$$

and $$\int_0^\infty\left(\frac{1}{1+t}-\frac{1}{1+t^2}\right)\frac{dt}{t}=\int_0^\infty\left(\frac{t}{1+t^2}-\frac{1}{1+t}\right)dt=0,$$

so that $$\int_0^\infty\left(\frac{1}{1+t}-e^{-t}\right)\frac{dt}{t}=\int_0^\infty\left(\frac{1}{1+t^2}-e^{-t}\right)\frac{dt}{t}.$$

Ex. 4. The two integrals

$$\int_0^\infty\left(\frac{e^{-t}}{t}-\frac{e^{-xt}}{e^t-1}\right)dt \quad\text{and}\quad \int_0^\infty\left(\frac{e^{-t}}{t}-\frac{1}{t(1+t)^{x+1}}\right)dt$$

are convergent and equal if $R(x+1)>0$.

When $t\to0$ the first integrand tends to $x-\frac{1}{2}$ and the second to x, so that the integrals converge at the lower limit. So far as convergence at ∞ is concerned the first integrand may be taken to be $e^{-(x+1)t}$ and the second $1/t(1+t)^{x+1}$ or simply $t^{-(x+2)}$, and therefore both integrals converge at infinity if $R(x+1)>0$.

Now the first integral may be considered as

$$\lim_{\lambda\to0}\int_\lambda^\infty\left(\frac{e^{-t}}{t}-\frac{e^{-xt}}{e^t-1}\right)dt,\ \lambda>0.$$

But, if $e^t-1=s$, we have

$$\int_\lambda^\infty\frac{e^{-xt}\,dt}{e^t-1}=\int_{e^\lambda-1}^\infty\frac{ds}{s(1+s)^{x+1}}=\int_\lambda^\infty\frac{ds}{s(1+s)^{x+1}}-\int_\lambda^{e^\lambda-1}\frac{ds}{s(1+s)^{x+1}}.$$

Let $x+1=\xi+i\eta$, where $\xi>0$; then

$$\int_\lambda^{e^\lambda-1}\frac{ds}{s(1+s)^\xi}$$

lies between $(1+\lambda)^{-\xi}\log[(e^\lambda-1)/\lambda]$ and $e^{-\lambda\xi}\log[(e^\lambda-1)/\lambda]$, and $\log[(e^\lambda-1)/\lambda]\to0$ when $\lambda\to0$, so that

$$\lim_{\lambda\to0}\int_\lambda^{e^\lambda-1}\frac{ds}{s(1+s)^{x+1}}=0.$$

Hence $$\lim_{\lambda\to0}\int_\lambda^\infty\left(\frac{e^{-t}}{t}-\frac{e^{-xt}}{e^t-1}\right)dt=\lim_{\lambda\to0}\int_\lambda^\infty\left(\frac{e^{-t}}{t}-\frac{1}{t(1+t)^{x+1}}\right)dt,$$

so that the two given integrals are equal.

Ex. 5. $\displaystyle\int_0^\infty \left\{(a-1)e^{-x}+\left(\frac{1}{x}+\frac{1}{2}\right)(e^{-ax}-e^{-x})\right\}\frac{dx}{x}=(a-\tfrac{1}{2})\log a-(a-1),$

where $a>0$, or, if a is complex, $R(a)>0$.

This integral is required in the next article. It may be derived from Ex. 11 in *Exercises* XIX by taking for the case (ii) $b=1$, $c=-(a-\frac{1}{2})$, but may easily be proved independently. Denote the integrand by $F(x)$; then, if $\lambda>0$,

$$\int_0^\infty F(x)dx=\mathop{L}_{\lambda\to 0}\int_\lambda^\infty F(x)dx.$$

Integrate $(e^{-ax}-e^{-x})/x^2$ by parts (the other terms of the integrand contain only the first power of x in the denominator); thus

$$\int_\lambda^\infty \frac{e^{-ax}-e^{-x}}{x^2}dx=\frac{e^{-a\lambda}-e^{-\lambda}}{\lambda}-\int_\lambda^\infty (ae^{-ax}-e^{-x})\frac{dx}{x},$$

and
$$\int_\lambda^\infty F(x)dx=\frac{e^{-a\lambda}-e^{-\lambda}}{\lambda}+(a-\tfrac{1}{2})\int_\lambda^\infty \frac{e^{-x}-e^{-ax}}{x}dx.$$

The given value follows at once since $(e^{-a\lambda}-e^{-\lambda})/\lambda\to-(a-1)$ when $\lambda\to 0$.

In the same way it may be proved that

$$\int_0^\infty\left\{\frac{e^{-2x}-e^{-x}}{x}+\tfrac{3}{2}e^{-2x}-\tfrac{1}{2}e^{-x}\right\}\frac{dx}{x}=-1+\tfrac{1}{2}\log 2.$$

This integral is also required in the next article.

167. Integral for $\log\Gamma(x)$. The integral will be derived from the expression for $\log\Gamma(x)$ as an infinite product; as in the preceding article the logarithms will be expressed by definite integrals all of which converge when $R(x)$ is positive.

Let $\displaystyle P_n(x)=\frac{n!\,n^{x-1}}{x(x+1)(x+2)\dots(x+n-1)}=n^{x-1}\prod_{r=1}^{n}\left(\frac{r}{x+r-1}\right);$

then

$$\log P_n(x)=(x-1)\log n-\sum_1^n\log\frac{x+r-1}{r}$$

$$=(x-1)\int_0^\infty\frac{e^{-t}-e^{-nt}}{t}dt-\sum_{r=1}^{n}\int_0^\infty \qquad\qquad \frac{\quad}{\quad}\,dt.$$

But
$$\sum_{r=1}^{n}\left[e^{-rt}-e^{-(x+r-1)t}\right]=\frac{(e^{-t}-e^{-xt})(1-e^{-nt})}{1-e^{-t}},$$

and therefore, taking together the terms that contain the factor e^{-nt}, we find

$$\log P_n(x)=\int_0^\infty\left\{(x-1)e^{-t}-\frac{e^{-t}-e^{-xt}}{1-e^{-t}}\right\}\frac{dt}{t}-R_n,$$

where
$$R_n=\int_0^\infty\left\{(x-1)-\frac{e^{-t}-e^{-xt}}{1-e^{-t}}\right\}\frac{e^{-nt}}{t}\cdot dt.$$

Now the coefficient of e^{-nt} in the integrand of R_n tends to $\frac{1}{2}x(x-1)$ when t tends to zero, and is finite when $t>0$, so that there is a finite number, K say, such that the integrand of R_n is, in absolute value, less than Ke^{-nt} for $t \geqq 0$. Hence

$$|R_n| < K\int_0^\infty e^{-nt}\,dt = \frac{K}{n},$$

and therefore $R_n \to 0$ when $n \to \infty$.

Again, the integrand of the other integral tends to $\frac{1}{2}(x-1)(x-2)$ when $t \to 0$, so that the integral converges at its lower limit; it is obviously convergent at ∞ if $R(x)>0$. Hence, if we let n tend to infinity, we find

$$\log\Gamma(x) = \int_0^\infty \left\{(x-1)e^{-t} - \frac{e^{-t}-e^{-xt}}{1-e^{-t}}\right\}\frac{dt}{t}. \quad \text{............(1)}$$

This integral may be expressed as the sum of two integrals, one of which can be evaluated in finite terms, while the other tends to zero when $R(x)$ tends to infinity. These integrals are

$$P(x) = \int_0^\infty \left\{(x-1)e^{-t} - \frac{e^{-t}}{1-e^{-t}} + \left(\frac{1}{t}+\frac{1}{2}\right)e^{-xt}\right\}\frac{dt}{t}, \quad \text{......(2)}$$

$$\mu(x) = \int_0^\infty \left\{\frac{1}{1-e^{-t}} - \frac{1}{t} - \frac{1}{2}\right\}\frac{e^{-xt}}{t}\,dt. \quad \text{......................(3)*}$$

Consider $\mu(x)$. The coefficient of e^{-xt} in the integrand is the function $f(t)/t$ of § 166, and therefore lies between 0 and 1/12; hence, if $x=\xi+i\eta$, $(\xi>0)$,

$$|\mu(x)| < \frac{1}{12}\int_0^\infty e^{-\xi t}\,dt = \frac{1}{12\xi}, \quad \text{.....................(4)}$$

and therefore $\mu(x) \to 0$ when $R(x) \to \infty$.

We give two methods of evaluating $P(x)$, the first of which holds whether x is real or complex, while the second assumes x to be real; both methods are somewhat artificial.

First Method; x complex. In (2) let $x=1$ and subtract $P(1)$ from $P(x)$, thus eliminating the term $e^{-t}/(1-e^{-t})$; then

$$P(x)-P(1) = \int_0^\infty \left\{(x-1)e^{-t} + \left(\frac{1}{t}+\frac{1}{2}\right)(e^{-xt}-e^{-t})\right\}\frac{dt}{t}$$
$$= (x-\tfrac{1}{2})\log x - (x-1)$$

by Ex. 5, § 166.

* $\mu(x)$ is Binet's notation for this function; Cauchy uses the notation $\varpi(x)$.

We have next to find $P(1)$, where

$$P(1)=\int_0^\infty\left\{\left(\frac{1}{t}+\frac{1}{2}\right)e^{-t}-\frac{1}{e^t-1}\right\}\frac{dt}{t}. \quad \ldots\ldots\ldots\ldots\ldots\ldots(a)$$

Change the variable from t to $2t$; then

$$P(1)=\int_0^\infty\left\{\frac{1}{2}\left(\frac{1}{t}+1\right)e^{-2t}-\frac{1}{e^{2t}-1}\right\}\frac{dt}{t}. \quad \ldots\ldots\ldots\ldots(b)$$

Next multiply equation (b) throughout by 2, and from the value $2P(1)$ given by (b) subtract the value $P(1)$ given by (a); we then find, after a slight reduction,

$$P(1)=\int_0^\infty\left\{\frac{e^{-2t}-e^{-t}}{t}+e^{-2t}-\tfrac{1}{2}e^{-t}+\frac{1}{e^t+1}\right\}\frac{dt}{t}$$

$$=\int_0^\infty\left\{\frac{1}{e^t+1}-\tfrac{1}{2}e^{-2t}\right\}\frac{dt}{t}+\int_0^\infty\left\{\frac{e^{-2t}-e^{-t}}{t}+\tfrac{3}{2}e^{-2t}-\tfrac{1}{2}e^{-t}\right\}\frac{dt}{t}.$$

The second of these integrals is, by Ex. 5 of § 166, equal to $(-1+\frac{1}{2}\log 2)$.

To find the other integral put $x=\frac{1}{2}$ in equation (1); then, since $\Gamma(\frac{1}{2})=\sqrt{\pi}$, we see that

$$\tfrac{1}{2}\log\pi=\int_0^\infty\left\{\frac{e^{-\frac{1}{2}t}-e^{-t}}{1-e^{-t}}-\tfrac{1}{2}e^{-t}\right\}\frac{dt}{t}=\int_0^\infty\left(\frac{1}{e^t+1}-\tfrac{1}{2}e^{-2t}\right)\frac{dt}{t}, \ldots(c)$$

by changing the variable from t to $2t$. Thus

$$P(1)=\tfrac{1}{2}\log\pi-1+\tfrac{1}{2}\log 2=-1+\log\sqrt{(2\pi)},$$

and $$P(x)=(x-\tfrac{1}{2})\log x-x+\log\sqrt{(2\pi)}. \quad \ldots\ldots\ldots\ldots\ldots(5)$$

Hence

$$\log\Gamma(x)=(x-\tfrac{1}{2})\log x-x+\log\sqrt{(2\pi)}+\mu(x)\ldots\ldots\ldots\ldots(6)$$

and, since

$$\log\Gamma(x+1)=\log\Gamma(x)+\log x,$$

$$\log\Gamma(x+1)=(x+\tfrac{1}{2})\log x-x+\log\sqrt{(2\pi)}+\mu(x). \quad \ldots\ldots\ldots(7)$$

Second Method; x real and positive. In equation (1) put y in place of x. The integral converges uniformly if

$$0<x\leqq y\leqq x+1,$$

and therefore we may integrate with respect to y under the integral sign. Hence

$$\int_x^{x+1}\log\Gamma(y)\,dy=\int_0^\infty\left\{(x-\tfrac{1}{2})e^{-t}-\frac{e^{-t}}{1-e^{-t}}+\frac{e^{-xt}}{t}\right\}\frac{dt}{t}.$$

Also $$\tfrac{1}{2}\log x = \tfrac{1}{2}\int_0^\infty \frac{e^{-t}-e^{-xt}}{t}\,dt,$$

so that $$P(x) = \int_x^{x+1} \log\Gamma(y)dy - \tfrac{1}{2}\log x$$
$$= (x-\tfrac{1}{2})\log x - x + \log\sqrt{(2\pi)}$$

by Ex. 12 of § 146. (In this method x must be real since integration with respect to a *complex variable* is not at our disposal.)

Ex. 1. Prove that, if $a>0$, $b>0$, $(b-a)P(1)$ is equal to
$$\int_0^\infty \left\{\frac{a}{e^{at}-1} - \frac{b}{e^{bt}-1} + \tfrac{1}{2}(be^{-bt}-ae^{-at}) + \frac{e^{-bt}-e^{-at}}{t}\right\}\frac{dt}{t},$$
and deduce that
$$\int_0^\infty \left\{\frac{a}{e^{at}-1} - \frac{b}{e^{bt}-1} + \tfrac{1}{2}(ae^{-at}-be^{-bt})\right\}\frac{dt}{t} = (b-a)\log\sqrt{(2\pi)}.$$

(Schlömilch.)

In the integral (a) for $P(1)$ change the variable to at then to bt and form the difference $bP(1)-aP(1)$, the factors a and b being taken with the transformed integrals in a and b respectively.

If $a=1$, $b=2$, the useful integral
$$\int_0^\infty \left[\frac{1}{e^t+1} + \tfrac{1}{2}(e^{-t}-2e^{-2t})\right]\frac{dt}{t} = \log\sqrt{(2\pi)}$$
is obtained.

Ex. 2. Prove Gauss's Formula
$$\Gamma(x)\Gamma\left(x+\frac{1}{n}\right)\Gamma\left(x+\frac{2}{n}\right)\dots\Gamma\left(x+\frac{n-1}{n}\right) = \frac{(2\pi)^{\frac{n-1}{2}}\Gamma(nx)}{n^{nx-\frac{1}{2}}}.$$

In the equation (1) take the n values x, $x+\dfrac{1}{n}$, $x+\dfrac{2}{n}, \dots, x+\dfrac{n-1}{n}$, and add the integrals; the logarithm of the product of the n functions is
$$\int_0^\infty \left\{\left(nx-\frac{n+1}{2}\right)e^{-t} - \frac{ne^{-t}}{1-e^{-t}} + \frac{e^{-t}}{1-e^{-t/n}}\right\}\frac{dt}{t}$$
$$= \int_0^\infty \left\{\left(nx-\frac{n+1}{2}\right)e^{-nt} - \frac{ne^{-nt}}{1-e^{-nt}} + \frac{e^{-nt}}{1-e^{-t}}\right\}\frac{dt}{t} \qquad \text{(i)}$$
by changing the variable to nt.

Now put nx for x in equation (1), and subtract the integral for $\log\Gamma(nx)$ from the integral (i); if P denote the product of the n functions $\Gamma(x)$, $\Gamma(x+1/n)$, ... we find, after a little reduction,
$$\log\left(\frac{P}{\Gamma(nx)}\right) = \int_0^\infty \left\{\frac{1}{e^t-1} - \frac{n}{e^{nt}-1} + \tfrac{1}{2}(e^{-t}-ne^{-nt})\right\}\frac{dt}{t}$$
$$+(\tfrac{1}{2}-nx)\int_0^\infty \frac{e^{-t}-e^{-nt}}{t}\,dt.$$

In Ex. 1 let $a=1$, $b=n$, and we obtain

$$\log\{P/\Gamma(nx)\}=\log\left\{(2\pi)^{\frac{n-1}{2}}\Big/ n^{nx-\frac{1}{2}}\right\},$$

and therefore
$$P=\frac{(2\pi)^{\frac{n-1}{2}}\Gamma(nx)}{n^{nx-\frac{1}{2}}}.$$

This proof of Gauss's formula is given by Schlömilch, *Compendium*, vol. II, p. 255.

Ex. 3. *Stirling's Approximation for* $n!$. Let n be a positive integer, and in equation (7) put n for x; then

$$\log(n!)=\log\{n^n e^{-n}\sqrt{(2\pi n)}e^{\mu(n)}\},$$

and therefore
$$n!=n^n e^{-n}\sqrt{(2\pi n)}e^{\mu(n)},\quad \mu(n)<\frac{1}{12n}.$$

The expression $(n/e)^n\sqrt{(2\pi n)}$ is called Stirling's Approximation for $n!$ when n is large, the *relative error* being less than $e^{1/12n}$; the *absolute error* may be "large" when the relative error is "small."

168. Asymptotic Expansion of $\log\Gamma(x)$. In equation (6) of the preceding article let x be real and positive; the term $\mu(x)$ may be expanded in powers of x^{-1} as follows:

$$\mu(x)=\int_0^\infty\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)\frac{e^{-xt}}{t}dt.$$

By § 94, equations (5) and (5*a*),

$$\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)\frac{1}{t}=\sum_{r=1}^{n}(-1)^{r-1}\frac{B_r}{(2r)!}t^{2r-2}+R_n'$$

where
$$R_n'=(-1)^n\theta_n'\frac{B_{n+1}}{(2n+2)!}t^{2n},\quad 0<\theta_n'<1.$$

Therefore, when the integrals are evaluated and expressions reduced,

$$\mu(x)=\sum_{r=1}^{n}(-1)^{r-1}\frac{B_r}{(2r-1)2r}\frac{1}{x^{2r-1}}+R_n(x),\quad \text{.........(1)}$$

where
$$R_n(x)=(-1)^n\frac{B_{n+1}}{(2n+2)!}\int_0^\infty\theta_n'e^{-xt}t^{2n}\,dt$$
$$=(-1)^n\theta_n\frac{B_{n+1}}{(2n+1)(2n+2)}\frac{1}{x^{2n+1}},\quad \text{..............(2)}$$

since
$$0<\int_0^\infty\theta_n'e^{-xt}t^{2n}\,dt<\int_0^\infty e^{-xt}t^{2n}\,dt=\frac{\Gamma(2n+1)}{x^{2n+1}},$$

so that $R_n(x)$ has the form stated, where $0<\theta_n<1$.

Let
$$u_r(x)=(-1)^{r-1}\frac{B_r}{(2r-1)2r}\frac{1}{x^{2r-1}},\quad S_n(x)=\sum_{r=1}^{n}u_r(x).$$

From the value of B_n in equation (6) of § 94 it is plain that

the infinite series $\Sigma u_r(x)$ is not convergent. On the other hand, if n and x are both fixed, the error $R_n(x)$ in taking $S_n(x)$ as the value of $\mu(x)$ is numerically less than $|u_{n+1}(x)|$, while, if n is fixed and x tends to infinity, $R_n(x)$ tends to zero. A non-convergent series of this type is called an *Asymptotic Series*, so that the function $\mu(x)$ is given by an *Asymptotic Series* or *Asymptotic Expansion*. For purposes of numerical calculation asymptotic expansions are very useful. See Bromwich, *Inf. Ser.* (2nd Ed.), Chapter XII.

Since $\log \Gamma(x) = P(x) + \mu(x)$, *the asymptotic expansion of* $\log \Gamma(x)$ is

$$(x-\tfrac{1}{2})\log x - x + \log\sqrt{(2\pi)} + \sum_{r=1}^{n}(-1)^{r-1}\frac{B_r}{(2r-1)2r}\frac{1}{x^{2r-1}}$$
$$+R_n(x) \quad \text{.........................(3)}$$

where $R_n(x)$ is given by (2).

Integral for $\mu(x)$. By Ex. 4 of § 165 the integral for $\mu(x)$ is

$$2\int_0^\infty \frac{\tan^{-1}(t/x)}{e^{2\pi t}-1}\,dt.$$

Ex. 1. $\displaystyle\int_0^\infty \frac{\tan^{-1}t\,dt}{e^{2\pi t}-1} = \tfrac{1}{2}\{1-\log\sqrt{(2\pi)}\}.$

In the equation $\log \Gamma(x) = P(x) + \mu(x)$ let $x = 1$.

Ex. 2. If $a > 0$ and $b > 0$ and $a/b = x$,

$$\log\{a(a+b)(a+2b)\dots(a+nb)\}$$
$$= \log a + n\log b + \log\Gamma(x+n+1) - \log\Gamma(x+1).$$

When x or a/b is large the factorial $a(a+b)\dots(a+nb)$ may be calculated by using the asymptotic expansion of $\log \Gamma(x)$.

169. Gauss's Function $\psi(x)$. This function is (§ 97) the derivative with respect to x of $\log\Gamma(x+1)$. Now, x being real and positive,

$$\Gamma(x+1) = x\Gamma(x) = \underset{n\to\infty}{L}\frac{n!\,n^x}{(x+1)(x+2)\dots(x+n)},$$

and therefore $$\psi(x) = \underset{n\to\infty}{L}\left\{\log n - \sum_{r=1}^{n}\frac{1}{x+r}\right\}. \quad \text{............(A)}$$

Further, $$\log n = \int_0^\infty \frac{e^{-t}-e^{-nt}}{t}\,dt, \quad \frac{1}{x+r} = \int_0^\infty e^{-(x+r)t}\,dt,$$

and, when these values are inserted in the expression for $\psi(x)$ and the terms that contain e^{-nt} separated from the rest, we find

$$\psi(x) = \int_0^\infty\left(\frac{e^{-t}}{t} - \frac{e^{-xt}}{e^t-1}\right)dt - \underset{n\to\infty}{L}\int_0^\infty\left(\frac{1}{t} - \frac{e^{-xt}}{e^t-1}\right)e^{-nt}\,dt.$$

The coefficient of e^{-nt} in the integrand of the second integral tends to $(x+\frac{1}{2})$ when t tends to zero, and is finite if t is positive; thus the coefficient of e^{-nt} is bounded, say less than K, if $t \geqq 0$, and therefore the integral is less than K/n, an expression which tends to zero when $n \to \infty$.

Again, the integrand of the first integral tends to $(x-\frac{1}{2})$ when t tends to zero, so that the integral converges at the lower limit; also it manifestly converges at infinity since x is positive.

Hence $$\psi(x)=\int_0^\infty\left(\frac{e^{-t}}{t}-\frac{e^{-xt}}{e^t-1}\right)dt \quad \text{.....................(1)}$$

$$=\int_0^\infty\left(\frac{e^{-t}}{t}-\frac{1}{t(1+t)^{x+1}}\right)dt \quad \text{...............(2)}$$

by Ex. 4 of § 166.

The expression (1) may also be obtained by using Weierstrass's form of $\Gamma(x+1)$ and using the value of γ given by Ex. 1 of § 166.

Another expression may be found by differentiating the equation (7) of § 167; thus

$$\psi(x)=\frac{d.\log\Gamma(x+1)}{dx}=\log x+\frac{1}{2x}+\frac{d.\mu(x)}{dx},$$

where $$\mu(x)=\int_0^\infty\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)\frac{e^{-xt}}{t}dt.$$

This integral may be differentiated with respect to x under the integral sign, since the integral obtained by differentiation converges uniformly in an arbitrarily large range $0<a\leqq x\leqq b$.

Hence $$\psi(x)=\log x+\frac{1}{2x}-\int_0^\infty\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)e^{-xt}dt$$

$$=\log x+\frac{1}{2x}-2\int_0^\infty\frac{t\,dt}{(x^2+t^2)(e^{2\pi t}-1)} \quad \text{.............(3)}$$

by Ex. 3 of § 165.

These expressions (1), (2) and (3) for $\psi(x)$ are valid if $R(x)>0$, but the proof given above does not show this, since differentiation with respect to a complex variable x lies outside our limits. If, however, $\psi(x)$ is taken to be defined by the limit (A), the values (1) and (2) hold for a complex x, since the various transformations depend only on a complex function of a real variable.

Ex. 1. If $R(x)>0$, $\dfrac{d\log\Gamma(x)}{dx}=\displaystyle\int_0^\infty\frac{e^{-t}-e^{-xt}}{1-e^{-t}}dt-\gamma$ ($\gamma=$Euler's Constant).

Differentiate equation (1) of § 167; thus

$$\frac{d \cdot \log \Gamma(x)}{dx} = \int_0^\infty \left(e^{-t} - \frac{t e^{-xt}}{1-e^{-t}}\right)\frac{dt}{t}$$

$$= \int_0^\infty \left(\frac{e^{-t} - e^{-xt}}{1-e^{-t}}\right) dt - \int_0^\infty \left(\frac{1}{1-e^{-t}} - \frac{1}{t}\right) e^{-t}\, dt,$$

and the last integral is equal to γ (Ex. 1, § 166).

Ex. 2. $\psi(x) + \gamma = \int_0^1 \frac{1-t^x}{1-t}\, dt, \quad x > 0.$

To the integral (1) for $\psi(x)$ add the integral for γ in Ex. 1 of § 166; this gives

$$\psi(x) + \gamma = \int_0^\infty \left(\frac{1-e^{-xt}}{1-e^{-t}}\right) e^{-t}\, dt = \int_0^1 \frac{1-s^x}{1-s}\, ds,$$

by changing the variable of integration to s, where $s = e^{-t}$.

When x is a (positive) rational number the above integral for $\psi(x) + \gamma$ can be expressed in terms of logarithms and circular functions; for example

$$\psi(\tfrac{2}{3}) + \gamma = \tfrac{3}{2} - \tfrac{3}{2}\log 3 + \frac{\pi}{2\sqrt{3}},$$

and the value of $\psi(n + \frac{2}{3}) + \gamma$, where n is a positive integer, can be expressed in terms of $\psi(\frac{2}{3}) + \gamma$ and rational fractions by formula (3) of § 97.

Ex. 3. $\psi(x) = -\int_0^1 \left(\frac{t^x}{1-t} + \frac{1}{\log t}\right) dt.$

Change the variable in the integral (1) from t to s, where $s = e^{-t}$.

170. Another Proof of the Integral for $\log \Gamma(x)$. The following proof, which is of frequent occurrence in the older textbooks, is merely sketched; it gives a good example of the tests for change of order in integration. The starting point is the integral for $\Gamma(x)$, where x is real and positive,

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\, dt.$$

The derivative $\Gamma'(x)$ may be obtained by differentiating under the integral sign (§ 158, Ex. 9), so that

$$\Gamma'(x) = \int_0^\infty e^{-t} t^{x-1} \log t\, dt = \int_0^\infty e^{-t} t^{x-1}\, dt \int_0^\infty \frac{e^{-s} - e^{-ts}}{s}\, ds,$$

by expressing $\log t$ as an integral. It may now be shown by Theorem III of § 162 that the order of integration may be changed; when the change has been made the integration with respect to t can be effected. Hence

$$\Gamma'(x) = \Gamma(x) \int_0^\infty \{e^{-s} - (1+s)^{-x}\} \frac{ds}{s} \qquad \text{.............(1)}$$

Now put y in place of x, and integrate $\Gamma'(y)/\Gamma(y)$ from $y=1$ to $y=x$; the integration under the integral sign is legitimate (§ 158, Ex. 7). Hence

$$\log \Gamma(x) = \int_0^\infty \left\{ e^{-s}.(x-1) - \frac{(1+s)^{-1}-(1+s)^{-x}}{\log(1+s)} \right\} \frac{ds}{s}.$$

In order to eliminate the term e^{-s}, put $x=2$; thus

$$0 = \int_0^\infty \left\{ e^{-s} - \frac{s(1+s)^{-2}}{\log(1+s)} \right\} \frac{ds}{s}.$$

Now multiply this equation by $(x-1)$ and subtract; therefore

$$\log \Gamma(x) = \int_0^\infty \left\{ \frac{x-1}{(1+s)^2} - \frac{(1+s)^{-1}-(1+s)^{-x}}{s} \right\} \frac{ds}{\log(1+s)} \quad \ldots(2)$$

and, if $1+s=e^t$, this becomes

$$\log \Gamma(x) = \int_0^\infty \left\{ (x-1)e^{-t} - \frac{e^{-t}-e^{-xt}}{1-e^{-t}} \right\} \frac{dt}{t}, \quad \ldots\ldots\ldots\ldots(3)$$

the same integral as (1) of § 167.

In equation (1) put $x+1$ in place of x; then

$$\psi(x) = \frac{\Gamma'(x+1)}{\Gamma(x+1)} = \int_0^\infty \left\{ e^{-s} - \frac{1}{(1+s)^{x+1}} \right\} \frac{ds}{s},$$

and this is the integral (2) of § 169.

Another method, due to Schaar and given by Hermite (*Cours lithographié*, 4th Ed. p. 128), may be sketched.

$$\frac{\Gamma(x)\Gamma(h)}{\Gamma(x+h)} = B(x, h) = \int_0^\infty \frac{t^{h-1}\,dt}{(1+t)^{x+h}}, \quad h>0.$$

Now
$$\frac{\Gamma(x+h)-\Gamma(x)}{h\Gamma(x+h)} = \frac{1}{h} - \frac{B(x,h)}{h\Gamma(h)} = \frac{\Gamma(h)-B(x,h)}{\Gamma(1+h)},$$

and
$$\Gamma(h) = \int_0^\infty e^{-t}t^{h-1}\,dt,$$

so that
$$\frac{\Gamma(x+h)-\Gamma(x)}{h\Gamma(x+h)} = \frac{1}{\Gamma(1+h)} \int_0^\infty \left\{ e^{-t} - \frac{1}{(1+t)^{x+h}} \right\} \frac{dt}{t^{1-h}}.$$

The limit for h tending to zero of the left-hand member of this equation is $\Gamma'(x)/\Gamma(x)$, and it has to be proved that the limit of the right-hand member is the integral

$$\int_0^\infty \left\{ e^{-t} - \frac{1}{(1+t)^x} \right\} \frac{dt}{t}.$$

The proof will form a good exercise.

171. Minimum Value of $\Gamma(x)$. The derivative of $\Gamma(x)$ (x real and positive) is given by the integral

$$\Gamma'(x) = \int_0^\infty e^{-t}t^{x-1}\log t\,dt = \int_1^\infty e^{-t}t^{x-1}\log t\,dt - \int_0^1 e^{-t}t^{x-1}\log\left(\frac{1}{t}\right)dt$$
$$= \varphi(x) - \psi(x), \text{ say.}$$

$\varphi(x)$ is a monotonic increasing function of x while $\psi(x)$ is a monotonic decreasing function of x, and therefore $\Gamma'(x)$ increases monotonically from $-\infty$ to ∞ as x increases from 0 to ∞. Hence $\Gamma'(x)$ vanishes once and only once as x increases from 0 to ∞, and changes from negative values when x is small to positive values when x is large, so that $\Gamma(x)$ has one minimum value.

Now $\Gamma(1)=\Gamma(2)$, and therefore the minimum value lies between $\Gamma(1)$ and $\Gamma(2)$. Calculation shows that the minimum occurs when x lies between 1·46 and 1·47, and $\Gamma(x)$ is just a little less than $\frac{1}{2}\sqrt{\pi}$; more accurately

$$x=1{\cdot}4616321\ldots \quad \Gamma(x)=0{\cdot}8856024\ldots,$$

when $\Gamma(x)$ is a minimum.

When negative values of x are admitted, as is the case when $\Gamma(x)$ is defined by the infinite product formula, there is an infinite number of negative values of x which make $\Gamma(x)$ a maximum or minimum, one and only one value of x lying in the interval $(-n, -n-1)$, where n is zero or a positive integer. The maxima are negative, and lie in the intervals $(0, -1)$, $(-2, -3)$, $(-4, -5)$, ...; the minima are positive, and lie in the intervals $(-1, -2)$, $(-3, -4)$, $(-5, -6)$, ... If n is large the value of x in the interval $(-n, -n+1)$ is very nearly $x=-n+(\log n)^{-1}$ when $\Gamma(x)$ is a maximum or minimum.

See Godefroy, *Théorie des Séries*, pp. 248-250, and the references there given. A graph of $\Gamma(x)$ will be found on p. 250 of Godefroy's book.

172. Integrals reducible to Gamma Functions. In this article a few examples will be given of integrals reducible to Gamma Functions; for an exhaustive treatment of the Gamma Function, with detailed indication of the sources of the various formulae, the student is referred to Nielsen's *Handbuch der Theorie der Gammafunktion*.

Many examples of integrals that are evaluated in terms of Gamma Functions have been already given in the text or among the Exercises.* The range of application is greatly extended by the use of the complex variable, but, even when the variable of integration is real, the Gamma Function

* See for example, Exercises XVI, XX.

provides for the evaluation of integrals of somewhat complicated character. For definiteness all constants are supposed to be real, unless they are expressly stated to be complex; in many cases, however, the student will have little difficulty in interpreting results for complex constants.

Type A. $\displaystyle\int_0^1 \frac{x^{m-1}(1-x)^{n-1}\,dx}{(a+bx)^{m+n+p}}$, a, $a+b$, m, n, positive.

Let $x/(a+bx)=t/(a+b)$, and the integral becomes

$$\frac{1}{a^{n+p}(a+b)^{m+p}}\int_0^1 t^{m-1}(1-t)^{n-1}(a+b-bt)^p\,dt.$$

If p is zero or a positive integer the integral is a sum of integrals each of which is a Beta Function and therefore expressible in terms of Gamma Functions; for other values of p the integral may be dealt with by use of a series.

If $x=\sin^2\theta$ the integral becomes

$$2\int_0^{\frac{\pi}{2}}\frac{(\sin\theta)^{2m-1}(\cos\theta)^{2n-1}d\theta}{\{a\cos^2\theta+(a+b)\sin^2\theta\}^{m+n+p}},$$

which therefore falls within the range of the above integral.

Another method is to express the given integral as a repeated integral by the substitution

$$\frac{\Gamma(m+n+p)}{(a+bx)^{m+n+p}}=\int_0^\infty e^{-(a+bx)y}\,y^{m+n+p-1}\,dy.$$

For examples see the Exercises on p. 456 of the *Elementary Treatise*.

Type B. Integrals derived from $B(m,n)$ by differentiating with respect to m or n, where $m>0$, $n>0$.

Here $$\int_0^1 x^{m-1}(1-x)^{n-1}\,dx=\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)},$$

and therefore, if the integral is differentiated with respect to m,

$$\int_0^1 x^{m-1}(1-x)^{n-1}\log x\,dx=\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}\{\psi(m-1)=\psi(m+n-1)\}, \quad\text{(i)}$$

since $$\frac{d\Gamma(x)}{dx}=\Gamma(x)\frac{d\log\Gamma(x)}{dx}=\Gamma(x)\,\psi(x-1).$$

If the values of $\psi(m-1)$ and $\psi(m+n-1)$ are expressed by means of the integral for $\psi(x)$ in § 169, Ex. 3, we may write the last equation in the form

$$\int_0^1 x^{m-1}(1-x)^{n-1}\log x\,dx=\int_0^1 x^{m-1}(1-x)^{n-1}\,dx\int_0^1\frac{t^{m+n-1}-t^{m-1}}{1-t}\,dt.$$

Again, differentiation with respect to n gives the integral

$$\int_0^1 x^{m-1}(1-x)^{n-1}\log(1-x)\,dx, \quad\ldots\ldots\ldots\ldots\ldots\ldots\text{(ii)}$$

and repeated differentiations with respect to m and n give integrals with integrands of the form

$$x^{m-1}(1-x)^{n-1}(\log x)^p(\log(1-x))^q,$$

when p and q are positive integers.

The change of variable from x to θ where $x=\cos^2\theta$ gives the equation

$$\int_0^{\frac{\pi}{2}}(\cos\theta)^{2m-1}(\sin\theta)^{2n-1}\,d\theta=\tfrac{1}{2}\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)},$$

and this form has been noticed in § 155, Ex. 1; simple examples are to be found in Exercises XVIII. Of course, for the evaluation of the integral (i) when m is a given number (such as $\frac{3}{4}$) the general formula is first calculated and then, when the differentiations have been effected, the particular value (such as $\psi(\frac{3}{4}-1)$) is taken.

In (ii) let $n=1$; then

$$\int_0^1 x^{m-1}\log(1-x)\,dx=\frac{\Gamma(m)\Gamma(1)}{\Gamma(m+1)}\{\psi(0)-\psi(m)\}=-\frac{\gamma+\psi(m)}{m},$$

a formula given by Abel.

Type C. In the Integral (1) of § 167 for $\log\Gamma(x)$ put in succession in place of x the numbers $a+b+1$, $a+1$ and $b+1$, where a, b, $(a+b)$ are each greater than -1, and express the sum

$$\log\Gamma(a+b+1)-\log\Gamma(a+1)-\log\Gamma(b+1)$$

as a single integral; the result is

$$\log\frac{\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(b+1)}=\int_0^\infty\frac{(1-e^{-at})(1-e^{-bt})}{1-e^{-t}}\frac{e^{-t}}{t}\,dt$$

$$=\int_0^1\frac{(1-x^a)(1-x^b)}{1-x}\frac{dx}{\log(1/x)}, \quad\ldots\ldots\text{(i)}$$

where $x=e^{-t}$.

In the same way it is proved that

$$\log\frac{\Gamma(a+1)\Gamma(a+b+c+1)}{\Gamma(a+b+1)\Gamma(a+c+1)}=\int_0^1\frac{x^a(1-x^b)(1-x^c)}{1-x}\frac{dx}{\log(1/x)}, \quad\text{(ii)}$$

where the argument of each Gamma Function is positive.

Since the logarithm on the left of (ii) is equal to

$$\log\frac{\Gamma(a+c+b+1)}{\Gamma(a+c+1)\Gamma(b+1)}-\log\frac{\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(b+1)},$$

the integral in (ii) may be obtained by putting $a+c$ in place of a in the integral (i) and then subtracting the integral (i) from the transformed integral.

Similarly in (i) put c for a, and take the difference of corresponding sides of (ii) and of the transformed (i); then

$$\log\frac{\Gamma(a+b+1)\Gamma(b+c+1)\Gamma(c+a+1)}{\Gamma(a+1)\Gamma(b+1)\Gamma(c+1)\Gamma(a+b+c+1)}$$
$$=\int_0^1\frac{(1-x^a)(1-x^b)(1-x^c)}{1-x}\frac{dx}{\log(1/x)}. \quad \text{...(iii)}$$

Type D. Dirichlet's and Liouville's Integral. In §§ 133, 157, integrals of the form

$$\int_0^k dx\int_0^{k-x}dy\int_0^{k-x-y}x^{m-1}y^{n-1}z^{p-1}f(x+y+z)\,dz \quad \text{.........}(a)$$

have been reduced by the change of variables

$$x+y+z=u,\quad x+y=uv,\quad x=uvw \quad \text{..............(i)}$$

to the expression

$$\frac{\Gamma(m)\Gamma(n)\Gamma(p)}{\Gamma(m+n+p)}\int_0^k u^{m+n+p-1}f(u)\,du \quad \text{..............}(b)$$

(It is, of course, understood that in this and the other examples the integrals are convergent.)

The field of integration may be defined as follows: the variables x, y, z are (i) never negative, and (ii) such that they satisfy the relation $\quad 0\leqq x+y+z\leqq k.$

The theorem expressed in the above transformation is quite general, as may be proved in the following way:

Let there be n variables $x_1, x_2, \dots, x_n$, and let

$$u_n=\iint\dots\int x_1^{\alpha_1-1}x_2^{\alpha_2-1}\dots x_n^{\alpha_n-1}f(x_1+x_2+\dots+x_n)\,dx_1\,dx_2\dots dx_n, \quad (c)$$

where the variables are (i) never negative, and (ii) such that they satisfy the relation

$$0\leqq x_1+x_2+\dots+x_n\leqq k;$$

then the integral u_n may be reduced to the form

$$u_n=\frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\dots\Gamma(\alpha_n)}{\Gamma(\alpha_1+\alpha_2+\dots+\alpha_n)}\int_0^k u^{\alpha_1+\alpha_2+\dots+\alpha_n-1}f(u)\,du. \quad \text{......}(d)$$

The theorem is true for $n=2$, $n=3$, the change of variables being of the type denoted by equations (i). Now apply the method of mathematical induction. Suppose the theorem to be true for the integral u_{n-1}, where

$$u_{n-1}=\iint\cdots\int x_2^{\alpha_2-1}x_3^{\alpha_3-1}\cdots x_n^{\alpha_n-1}$$
$$\times f(x_1+x_2+\ldots+x_n)\,dx_2\,dx_3\ldots dx_n, \quad\ldots\ldots\ldots\ldots(e)$$

the variables $x_2, x_3, \ldots, x_n$ being never negative and satisfying the relation $\quad 0\leqq x_2+x_3+\ldots+x_n\leqq k-x_1.$

By hypothesis u_{n-1} can be transformed so as to become

$$u_{n-1}=\frac{\Gamma(\alpha_2)\Gamma(\alpha_3)\ldots\Gamma(\alpha_n)}{\Gamma(\alpha_2+\alpha_3+\ldots+\alpha_n)}\int_0^{k-x_1}y_2^{\alpha_2+\alpha_3+\ldots+\alpha_n-1}f(x_1+y_2)\,dy_2 \quad (f)$$

by the change of variables given by the equations

$$x_2+x_3+\ldots+x_n=y_2,\; x_2+x_3+\ldots+x_{n-1}=y_2y_3,$$
$$x_2+x_3+\ldots+x_{n-2}=y_2y_3y_4, \ldots, x_2=y_2y_3\ldots y_n.$$

If the coefficient of the integral in (f) is denoted by A, and if $\alpha_2+\alpha_3+\ldots+\alpha_n=\beta$, the integral u_n will therefore be

$$u_n=A\int_0^k x_1^{\alpha_1-1}\,dx_1\int_0^{k-x_1}y_2^{\beta-1}f(x_1+y_2)\,dy_2.$$

Now let $x_1+y_2=z_1$, $x_1=z_1z_2$, and we find

$$u_n=A\,\frac{\Gamma(\alpha_1)\Gamma(\beta)}{\Gamma(\alpha_1+\beta)}\int_0^k z_1^{\alpha_1+\beta-1}f(z_1)\,dz.$$

If A is given its value this equation is simply equation (d), with z_1 instead of u as the variable of integration.

The change of variables in passing from the form (c) to the form (d) is, if $z_2'=1-z_2$,

$$x_2+x_3+\ldots+x_n+x_1=z_1,\; x_2+x_3+\ldots+x_n=z_1z_2',$$
$$x_2+x_3+\ldots+x_{n-1}=z_1z_2'y_3, \ldots, x_2=z_1z_2'y_3\ldots y_n,$$

and these are of the type (i).

If, in the case of three variables, the function f is not $f(x+y+z)$ but

$$f\left\{\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2+\left(\frac{z}{c}\right)^2\right\}$$

and the field of integration is the region bounded by the ellipsoid

$$x^2/a^2+y^2/b^2+z^2/c^2=1,\; x\geqq 0,\; y\geqq 0,\; z\geqq 0,$$

let $(x/a)^2=\xi$, $(y/b)^2=\eta$, $(z/c)^2=\zeta$, and change the variables to

ξ, η, ζ. The new field is given by the relation $0 \leqq \xi + \eta + \zeta \leqq 1$, so that the integral falls under the type just discussed. Obviously a similar change of variables is effective if the index of $x/a, \ldots$ is n instead of 2, and in many other cases. See, for example, Exercises XVI, 15, 24.

Type E. $$\int_0^k dx \int_0^{k-x} \frac{x^{m-1}y^{n-1}\,dy}{(\alpha + ax + by)^{m+n}},$$

where α, k, m, n are positive and a, b positive or zero.

First change the variables to u, v, where $x + y = u$, $x = uv$; the integral becomes

$$\int_0^k u^{m+n-1}\,du \int_0^1 \frac{v^{m-1}(1-v)^{n-1}\,dv}{[\alpha + bu + (a-b)uv]^{m+n}}.$$

Next let $\alpha + bu = A$, $(a-b)u = B$, so that $A + B = \alpha + au > 0$, and apply the substitution (Type A above), namely

$$v/(A + Bv) = \xi/(A + B);$$

the integral with respect to v is equal to

$$\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} \frac{1}{A^n(A+B)^m},$$

so that the given integral becomes

$$\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} \int_0^k \frac{u^{m+n-1}\,du}{(\alpha + au)^m(\alpha + bu)^n}.$$

If the index of $(\alpha + ax + by)$ is $m + n + p$ the integral with respect to v will be

$$\frac{1}{A^{n+p}(A+B)^{m+p}} \int_0^1 \xi^{m-1}(1-\xi)^{n-1}(A + B - B\xi)^p\,d\xi.$$

Hence the given repeated integral may be expressed when p is a positive integer (or zero) as a sum of simple integrals with respect to u.

A repeated integral in three or more variables may be reduced in the same way; thus the integral for three variables is

$$\int_0^k dx \int_0^{k-x} dy \int_0^{k-x-y} \frac{x^{m-1}y^{n-1}z^{p-1}\,dz}{(\alpha + ax + by + cz)^{m+n+p}},$$

where α, k, m, n, p are positive and a, b, c positive or zero.

First apply the transformation $x + y + z = u$, $x + y = uv$, $x = uvw$, and the integral becomes

$$\int_0^k u^{m+n+p-1}\,du \int_0^1 \int_0^1 \frac{v^{m+n-1}(1-v)^{p-1}\,w^{m-1}(1-w)^{n-1}\,dv\,dw}{[\alpha + cu + (b-c)uv + (a-b)uvw]^{m+n+p}}.$$

Next take the integral with respect to v (which contains all the indices m, n, p) and let

$$\alpha + cu = A,\quad (b-c)u + (a-b)uw = B;$$

then
$$A + B = \alpha + bu + (a-b)uw > 0.$$

The substitution $v/(A+Bv) = \xi/(A+B)$ gives as the value of the integral with respect to v,

$$\frac{\Gamma(m+n)\Gamma(p)}{\Gamma(m+n+p)} \cdot \frac{1}{(\alpha+cu)^p} \cdot \frac{1}{[\alpha+bu+(a-b)uw]^{m+n}}.$$

The integral with respect to w is now of the Type A, and finally the given integral is seen to be equal to

$$\frac{\Gamma(m)\Gamma(n)\Gamma(p)}{\Gamma(m+n+p)} \int_0^k \frac{u^{m+n+p-1}\,du}{(\alpha+au)^m(\alpha+bu)^n(\alpha+cu)^p}.$$

In the same way it may be proved that when the integrand contains the factor $f(x+y+z)$ the transformed integral contains the factor $f(u)$.

EXERCISES XXII.

The examples in this set of Exercises are well-known theorems or very obvious deductions from such theorems; for information on the sources of these theorems the student is referred to Nielsen's *Handbuch*.

1. $\log \Gamma(x) = \int_0^1 \left\{\frac{1-t^{x-1}}{1-t} - (x-1)\right\} \frac{dt}{\log t},\quad R(x) > 0.$

2. $\frac{d \,.\log \Gamma(x)}{dx} = \int_0^1 \frac{t^{x-1}-1}{t-1}\,dt - \gamma,\quad R(x) > 0.$

3. $\frac{d\,.\log \Gamma(x)}{dx} = \int_0^\infty \left(\frac{e^{-t}}{t} - \frac{e^{-xt}}{1-e^{-t}}\right) dt,\quad R(x) > 0$

$$= \log x - \int_0^1 \left(\frac{1}{1-t} + \frac{1}{\log t}\right) t^{x-1}\,dt.$$

4. If $0 < R(x) < 1$,

(i) $n\gamma + \sum_{r=0}^{n-1} \psi\left(x + \frac{r}{n}\right) = \int_0^1 \left(\frac{n}{1-t} - \frac{t^x}{1-t^{1/n}}\right) dt;$

(ii) $\gamma + \frac{1}{n}\sum_{r=0}^{n-1} \psi\left(x + \frac{r}{n}\right) = \int_0^1 \left(\frac{n}{1-t^n} - \frac{t^{nx}}{1-t}\right) t^{n-1}\,dt.$

5. If m and n are positive integers with no common factor and m less than n, show from Ex. 2 that

$$\psi\left(\frac{m}{n} - 1\right) + \gamma = n\int_0^1 \frac{x^{m-1} - x^{n-1}}{x^n - 1}\,dx,$$

and, by integrating the fraction, prove that

$$\psi\left(\frac{m}{n}-1\right)+\gamma=-\log n-\frac{\pi}{2}\cot\frac{m\pi}{n}+\sum_{r=1}^{n-1}\cos\left(\frac{2rm\pi}{n}\right)\log\left(2\sin\frac{r\pi}{n}\right).$$

[This method of proving Gauss's Formula (Gauss's *Werke*, III, p. 157) is laborious; for other methods see Nielsen, p. 20, or Bromwich, *Inf. Series* (2nd Ed.), p. 522, Ex. 43.]

6. $2\log\Gamma(x)+\log\left(\dfrac{\sin\pi x}{\pi}\right)=\log\Gamma(x)-\log\Gamma(1-x)$

$$=\int_0^\infty\left\{\frac{e^{(\frac{1}{2}-x)t}-e^{-(\frac{1}{2}-x)t}}{e^{\frac{1}{2}t}-e^{-\frac{1}{2}t}}-(1-2x)e^{-t}\right\}\frac{dt}{t},\ 0<x<1.$$

7. If $0<x<1$,

(i) $\dfrac{e^{(\frac{1}{2}-x)t}-e^{-(\frac{1}{2}-x)t}}{e^{\frac{1}{2}t}-e^{-\frac{1}{2}t}}=4\displaystyle\sum_{n=1}^{\infty}\frac{2n\pi\sin 2n\pi x}{4n^2\pi^2+t^2}$;

(ii) $1-2x=4\displaystyle\sum_{n=1}^{\infty}\frac{\sin 2n\pi x}{2n\pi}$;

(iii) $\displaystyle\int_0^\infty\left\{\frac{2n\pi}{4n^2\pi^2+t^2}-\frac{e^{-t}}{2n\pi}\right\}\frac{dt}{t}=\frac{\gamma+\log 2\pi+\log n}{2n\pi}$,

where γ is Euler's Constant.

8. Deduce from Examples 6 and 7 that, if $0<x<1$,

(i) $2\log\Gamma(x)+\log\left(\dfrac{\sin\pi x}{\pi}\right)=4\displaystyle\sum_{n=1}^{\infty}\frac{\gamma+\log 2\pi+\log n}{2n\pi}\sin 2n\pi x$;

(ii)
$$\begin{aligned}\log\Gamma(x)&=(\tfrac{1}{2}-x)(\gamma+\log 2)+(1-x)\log\pi-\tfrac{1}{2}\log\sin\pi x\\&\qquad+\sum_{n=1}^{\infty}\frac{\log n}{n\pi}\sin 2n\pi x\\&=(\tfrac{1}{2}-x)\gamma+(1-x)\log\pi-\tfrac{1}{2}\log\sin\pi x\\&\qquad+\sum_{n=1}^{\infty}\frac{\log 2n}{n\pi}\sin 2n\pi x.\end{aligned}$$

9. If $0<x<1$, show that

$$\log\sin\pi x=-\log 2-\sum_{n=1}^{\infty}\frac{\cos 2n\pi x}{n},$$

and deduce that

(i) $\displaystyle\int_0^1\log(\sin\pi x)\,.\,\sin 2n\pi x\,dx=0$;

(ii) $\displaystyle\int_0^1\log(\sin\pi x)\,.\,\cos n\pi x\,dx=\begin{cases}-1/n,\ n\text{ even}\\ 0,\ n\text{ odd.}\end{cases}$

10. Prove the following relations:

(i) $\displaystyle\int_0^1\log\Gamma(x)\,.\,\sin 2n\pi x\,dx=\frac{\gamma+\log 2n\pi}{2n\pi}$;

(ii) $\displaystyle\int_0^1\log\Gamma(x)\,.\,\cos 2n\pi x\,dx=\frac{1}{4n}$.

For developments where $\psi(x)$ takes the place of $\log\Gamma(x)$ see Nielsen, *Handbuch*, pp. 202-204. For example, if $0<x<1$,

11. $\psi(x-1)\sin\pi x + \frac{\pi}{2}\cos\pi x + (\gamma + \log 2\pi)\sin\pi x$

$$= \sum_{n=1}^{\infty} \log\left(\frac{n}{n+1}\right)\sin(2n+1)\pi x.$$

12. If $\beta(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{x+n}$, $x > 0$, prove that

(i) $\beta(x) = \int_0^1 \frac{t^{x-1}dt}{1+t} = \int_0^{\infty} \frac{e^{-xt}dt}{1+e^{-t}}$;

(ii) $\beta(x) = -\frac{d}{dx}\log B\left(\frac{x}{2}, \frac{1}{2}\right) = -\frac{d}{dx}\log\frac{\Gamma\left(\frac{x}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{x+1}{2}\right)}$

(iii) $\frac{x-1}{4}\,\beta\left(\frac{x-1}{2}\right) = \frac{1}{4} + \int_0^1 \frac{t^x\,dt}{(1+t^2)^2}$, $x > 1$.

13. Prove the following properties of $\beta(x)$:

(i) $\beta(x) + \beta(1+x) = \frac{1}{x}$; (ii) $\beta(x) + \beta(1-x) = \frac{\pi}{\sin\pi x}$;

(iii) $\beta(x) = \frac{1}{2}\left\{\psi\left(\frac{x-1}{2}\right) - \psi\left(\frac{x-2}{2}\right)\right\} = \psi(x) - \psi\left(\frac{x}{2}\right) + \frac{1}{x} - \log 2$;

(iv) $\beta(1) = \log 2$; (v) $\beta(\frac{1}{2}) = \frac{\pi}{2}$.

14. Deduce from Ex. 12 that

(i) $\int_0^{\infty} \frac{e^{-t} - e^{-xt}}{1+e^{-t}}\frac{dt}{t} = \log\pi - \log B\left(\frac{x}{2}, \frac{1}{2}\right)$;

(ii) $\int_0^{\infty}\left(\frac{1}{2} - \frac{1}{e^t+1}\right)\frac{e^{-t}}{t}\,dt = \frac{1}{2}\log\frac{\pi}{2}$.

15. If $\nu(x) = \sum_{n=0}^{\infty}\left\{\frac{1}{x+n} - \log\left(1 + \frac{1}{x+n}\right)\right\}$, $R(x) > 0$, prove that

$$\nu(x) = \int_0^{\infty}\left(\frac{1}{e^t-1} - \frac{1}{t} + 1\right)e^{-xt}\,dt$$

$$= \frac{1}{2x} + 2\int_0^{\infty}\frac{t\,dt}{(x^2+t^2)(e^{2\pi t}-1)}$$

Deduce the relations:

(i) $\nu(x) + \psi(x) = \log x + \frac{1}{x}$; (ii) $\nu(x) + \frac{d\,.\,\mu(x)}{dx} = \frac{1}{2x}$;

(iii) $\frac{d\,.\,\mu(x)}{dx} = \psi(x) - \log x - \frac{1}{2x}$; (iv) $\nu(1) = \gamma$;

(v) $\nu(\frac{1}{2}) = \gamma + \log 2$.

16. Verify that

$$\mu(x) = \sum_{n=0}^{\infty}\left\{(x+n+\tfrac{1}{2})\log\left(1+\frac{1}{x+n}\right) - 1\right\}.$$

$\not{d}$t

$$\tfrac{1}{2}\int_0^1 (t-t^2)\frac{d\psi(t+x-1)}{dt}\,dt;$$

$$(x)=\frac{1}{2x}-\tfrac{1}{2}\int_0^1 (t-t^2)\frac{d^2\psi(t+x-1)}{dt^2}\,dt.$$

rove that

(i) $\mu(x)-\mu(2x)=2 \qquad \dfrac{\tan^{-1}\left(\dfrac{t}{x}\right)dt}{e^{2\pi t}+1}$.

(ii) $\nu(x)-2\nu(2x)=2\displaystyle\int_0^\infty \frac{t\,dt}{(x^2+t^2)(e^{2\pi t}+1)};$

(iii) $\qquad \beta(x+\tfrac{1}{2})=2\displaystyle\int_0^\infty \frac{x\,dt}{(x^2+t^2)(e^{\pi t}+e^{-\pi t})}.$

19. $\displaystyle\int_0^\infty e^{-yt}t^{x-1}\log t\,dt=\frac{\Gamma(x)}{y^x}\left\{\psi(x-1)-\log y\right\},\quad \begin{matrix}x>0,\\ y>0.\end{matrix}$

20. $\Gamma(x)=\displaystyle\int_0^\infty e^{-t}(t-x)t^{x-1}\log t\,dt,\ x>0.$

21. $\dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}=\dfrac{1}{x}-\dfrac{y-1}{x+1}+\dfrac{(y-1)(y-2)}{2!\,(x+2)}-\dfrac{(y-1)(y-2)(y-3)}{3!\,(x+3)}+\dots,\quad \begin{matrix}x>0,\\ y>0.\end{matrix}$

22. By using the expression for $F(\alpha,\beta,\gamma,1)$ in terms of Gamma Functions, prove that, if $x>0$,

$$\left\{\frac{\Gamma(x)}{\Gamma(x+\frac{1}{2})}\right\}^2=\frac{1}{x}+\sum_{r=1}^\infty \frac{1^2.3^2.5^2\dots(2r-1)^2}{4.8.12\dots4r}\,\frac{1}{x(x+1)(x+2)\dots(x+r)}.$$

23. If n and p are positive integers and $x>0$, show that

$$\sum_{r=0}^{n-1}\log\Gamma\left(px+\frac{rp}{n}\right)=\int_0^\infty f(t)\frac{dt}{t}=\int_0^\infty f(nt)\frac{dt}{t},$$

where $f(t)$ is the function

$$[n(px-1)+\tfrac{1}{2}p(n-1)]e^{-t}-\frac{ne^{-t}}{1-e^{-t}}+\frac{e^{-pxt}(1-e^{-pt})}{(1-e^{-t})(1-e^{-pt/n})},$$

so that $f(nt)$ may be expressed in the form

$$\left[npx+\frac{np-n-p}{2}-\frac{n}{2}\right]e^{-nt}-\frac{n}{e^{nt}-1}+\frac{e^{-npxt}(1-e^{-npt})}{(1-e^{-nt})(1-e^{-pt})}.$$

24. In Ex. 23 interchange n and p, and take the difference of the two sums; then

$$\sum_{r=0}^{n-1}\log\Gamma\left(px+\frac{rp}{n}\right)-\sum_{r=0}^{p-1}\log\Gamma\left(nx+\frac{rn}{p}\right)$$

$$=\left(npx+\frac{np-n-p}{2}\right)\int_0^\infty \frac{e^{-nt}-e^{-pt}}{t}\,dt$$

$$+\int_0^\infty\left\{\frac{p}{e^{pt}-1}-\frac{n}{e^{nt}-1}+\tfrac{1}{2}\left(pe^{-pt}-ne^{-nt}\right)\right\}\frac{dt}{t}$$

$$=\left(npx+\frac{np-n-p}{2}\right)\log\frac{p}{n}+(n-p)\log\sqrt{(2\pi)}.$$

(See § 167, Ex. 1.)

25. Deduce from Ex. 24 that

$$\prod_{r=0}^{n-1}\Gamma\left(px+\frac{rp}{n}\right)=(2\pi)^{\frac{n-p}{2}}\left(\frac{p}{n}\right)^{npx+\frac{np-n-p}{2}}\prod_{r=0}^{p-1}\Gamma\left(nx+\frac{rn}{p}\right).$$

This generalisation of Gauss's Multiplication Formula is due to Schobloch (Nielsen, *l.c.* p. 198); $p=1$ gives Gauss's Formula.

EXERCISES XXIII.

1. $\displaystyle\int_0^1 \frac{x^2(1-x)^3\,dx}{(1+x)^6}=\frac{1}{768}.$

2. $\displaystyle\int_0^1 \frac{\sqrt{(1+x^2)}}{\sqrt{(1-x^2)}}x^2dx=\frac{1}{12\sqrt{(2\pi)}}\{[\Gamma(\tfrac{1}{4})]^2+12[\Gamma(\tfrac{3}{4})]^2\}.$

3. $\displaystyle\int_0 \frac{dx}{\left(1+x^{\frac{1}{n}}\right)^{2n}}=\frac{\Gamma(n+1)\Gamma(n+1)}{\Gamma(2n+1)}.$

4. $\displaystyle\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(1+\sin x)^m(1-\sin x)^n\,dx=2^{m+n}B(m+\tfrac{1}{2},\,n+\tfrac{1}{2}).$

5. $\displaystyle\int_0^{\frac{\pi}{2}}\cos^{2m-1}x\log(\cos x)dx=\frac{\sqrt{\pi}}{4}\frac{\Gamma(m)}{\Gamma(m+\frac{1}{2})}\{\psi(m-1)-\psi(m-\tfrac{1}{2})\}.$

6. In the integral $\displaystyle\int_0^1 \frac{x^a(1-x^b)(1-x^c)}{1-x}\frac{dx}{\log(1/x)}$, let $x=t^2$ and prove that, if $a+1>0$, $a+b+1>0$,

$$\int_0^1 \frac{t^{2a+1}(1-t^{2b})dt}{(1+t)\log t}=\log\frac{\Gamma(a+\frac{3}{2})\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(a+b+\frac{3}{2})}.$$

7. Deduce from Ex. 6 that, if $\alpha>0$, $\beta>0$,

$$\int_0^1 \frac{x^{\alpha-1}-x^{\beta-1}}{(1+x)\log x}dx=\log\frac{\Gamma\left(\frac{\alpha+1}{2}\right)\Gamma\left(\frac{\beta}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)\Gamma\left(\frac{\beta+1}{2}\right)},$$

and that, if $c>0$,

$$\int_0^\infty e^{-ct}\tanh t\frac{dt}{t}=\log\frac{c}{4}+2\log\left\{\Gamma\left(\frac{c}{4}\right)\Big/\Gamma\left(\frac{c+2}{4}\right)\right\}.$$

8. If the argument of each Gamma Function is positive, show that

$$\int_0^1 \frac{x^a(1-x^b)(1-x^c)(1-x^d)}{1-x}\frac{dx}{\log(1/x)}$$

$$=\log\frac{\Gamma(a+1)\Gamma(a+b+c+1)\Gamma(a+b+d+1)\Gamma(a+c+d+1)}{\Gamma(a+b+1)\Gamma(a+c+1)\Gamma(a+d+1)\Gamma(a+b+c+d+1)}.$$

Deduce that

$$\int_0^1 \frac{t^{2a+1}(1-t^{2b})^2dt}{(1+t)\log t}=\log\frac{\Gamma(a+\frac{3}{2})\Gamma(a+2b+\frac{3}{2})[\Gamma(a+b+1)]^2}{\Gamma(a+1)\Gamma(a+2b+1)[\Gamma(a+b+\frac{3}{2})]^2}.$$

9. From Ex. 8, or otherwise, show that if $c>0$,

$$\int_0^\infty e^{-ct}(1-\operatorname{sech} t)\frac{dt}{t}=-\log\frac{c}{4}+2\log\left\{\Gamma\left(\frac{c+3}{4}\right)\Big/\Gamma\left(\frac{c+1}{4}\right)\right\}.$$

10. If $a+1$ is positive, show that

$$\int_0^1 \frac{x^a(x-1)^2}{\log x}\,dx=\log(a+3)-2\log(a+2)+\log(a+1),$$

and, generally, if n is a positive integer,

$$\int_0^1 \frac{x^a(x-1)^n}{\log x}\,dx=(-1)^{n-1}\int_0^\infty e^{-(a+1)t}(1-e^{-t})^n\frac{dt}{t}$$

$$=\sum_{r=0}^{n}(-1)^r{}_nC_r\log(n+a+1-r).$$

11. $\displaystyle\int_0^1 \frac{x^{a-1}dx}{1+x}=\tfrac{1}{2}\left\{\psi\left(\frac{a-1}{2}\right)-\psi\left(\frac{a-2}{2}\right)\right\},\ a>0.$

[See *Exercises* XXII, 12, 13.]

Deduce that

$$\int_0^\infty \frac{e^{-ax}dx}{\cosh x}=\tfrac{1}{2}\left\{\psi\left(\frac{a-1}{4}\right)-\psi\left(\frac{a-3}{4}\right)\right\}.$$

12. Prove that, if $-1<\alpha<1$,

$$\int_0^\infty \frac{\sinh\alpha x}{\cosh x}\frac{dx}{x}=\int_0^1 \frac{t^{-\frac{1+\alpha}{2}}(1-t^\alpha)}{(1+t)\log(1/t)}\,dt,$$

and apply Ex. 6 to prove that the integral is equal to

$$\log\left\{\frac{\Gamma\left(\frac{1-\alpha}{4}\right)\Gamma\left(\frac{3+\alpha}{4}\right)}{\Gamma\left(\frac{1+\alpha}{4}\right)\Gamma\left(\frac{3-\alpha}{4}\right)}\right\},\ \text{that is, } \log\cot\left(\frac{1-\alpha}{4}\pi\right).$$

Deduce by differentiation of the above integral that

$$\int_0^\infty \frac{\cosh\alpha x}{\sinh x}\,dx=\frac{\pi}{2}\sec\frac{\pi\alpha}{2},\quad 0<\alpha<1.$$

13. If a and b are positive and the integral

$$\int_0^\infty F\left\{\left(ax-\frac{b}{x}\right)^2\right\}dx$$

convergent, prove that the integral is equal to

$$\frac{1}{a}\int_0^\infty F(x^2)\,dx.$$

If a, b, $(a+b-c)$ are all positive and $n+\frac{1}{2}>0$, show, after changing the variable from x to t, where $x=t^2/(1+t^2)$, that

$$\int_0^1 \frac{x^{n+\frac{1}{2}}(1-x)^{1-\frac{1}{2}}dx}{(a+bx-cx^2)^{n+1}}=\frac{\sqrt{\pi}\,\Gamma(n+\frac{1}{2})}{\beta\{c+(\alpha+\beta)^2\}^{n+\frac{1}{2}}\Gamma(n+1)},$$

where $\alpha=|a^{\frac{1}{2}}|$, $\beta=|(a+b-c)^{\frac{1}{2}}|$.

14. Prove that

(i) $\int_0^1 \frac{dx}{\sqrt{(1-x^4)}} = \frac{[\Gamma(\frac{1}{4})]^2}{4\sqrt{(2\pi)}}$; (ii) $\int_0^1 \frac{x^2dx}{\sqrt{(1-x^4)}} = \frac{[\Gamma(\frac{3}{4})]^2}{\sqrt{(2\pi)}}$;

(iii) $\int_0^1 \frac{dx}{\sqrt{(1+x^4)}} = \frac{[\Gamma(\frac{1}{4})]^2}{8\sqrt{\pi}}$.

15. If α, a, b, m, n, p are all positive, show that

$$\int_0^1 dx \int_0^{1-x} \frac{x^{m-1}y^{n-1}(1-x-y)^{p-1}}{(\alpha+ax+by)^{m+n+p+1}}\,dy$$
$$=\frac{\Gamma(m)\Gamma(n)\Gamma(p)}{\Gamma(m+n+p+1)}\left(\frac{m}{\alpha+a}+\frac{n}{\alpha+b}+\frac{p}{\alpha}\right)\frac{1}{(\alpha+a)^m(\alpha+b)^n\alpha^p}.$$

16. If the constants are all positive, prove that

$$\int_0^\infty \int_0^\infty \frac{e^{-(ax+by)}x^{m-1}y^{n-1}}{(\alpha+\beta x+\gamma y)^p}\,dx\,dy$$
$$=\frac{1}{\Gamma(p)}\int_0^\infty\int_0^\infty\int_0^\infty e^{-(ax+by)}e^{-(\alpha+\beta x+\gamma y)z}x^{m-1}y^{n-1}z^{p-1}\,dx\,dy\,dz$$
$$=\frac{\Gamma(m)\Gamma(n)}{\Gamma(p)}\int_0^\infty \frac{e^{-\alpha z}z^{p-1}dz}{(a+\beta z)^m(b+\gamma z)^n}.$$

Extend the theorem to the case of n variables.

17. The density at the point (x, y, z) of the solid bounded by the ellipsoid $x^2/a^2+y^2/b^2+z^2/c^2=1$ is

$$\mu\left\{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}\right\}^{n-1},$$

where μ is constant and n is positive. Prove that the mean density of the solid is $3\mu 2^{2n-2}\{\Gamma(n)\}^2/(2n+1)\Gamma(2n)$.

18. If the variables $x_1, x_2, \dots x_n$ are such that

$$0 \leqq x_1^2+x_2^2+\dots+x_n^2 \leqq 1,$$

prove that the variables $x_1, x_2, \dots x_n$ may be changed to new variables $\xi_1, \xi_2, \dots \xi_n$ (see § 134, Ex. 5) so that the integral

$$\iint\dots\int F(a_1x_1+a_2x_2+\dots+a_nx_n)\,dx_1dx_2\dots dx_n$$

shall become $$\iint\dots\int F(k\xi_1)\,d\xi_1d\xi_2\dots d\xi_n,$$

where $k=|(a_1^2+a_2^2+\dots+a_n^2)^{\frac{1}{2}}|$ and $0\leqq\xi_1^2+\xi_2^2+\dots+\xi_n^2\leqq 1$ (ξ_r positive), and then show that the integral is equal to

$$\frac{[\Gamma(\frac{1}{2})]^{n-1}}{\Gamma\left(\frac{n-1}{2}+1\right)}\int_{-1}^1 F(k\xi_1)(1-\xi_1^2)^{\frac{n-1}{2}}\,d\xi_1.$$

19. If in Ex. 18 the integrand is

$$F(a_1x_1+a_2x_2+\dots+a_nx_n)/\sqrt{(1-x_1^2-x_2^2-\dots-x_n^2)},$$

prove that the integral becomes

$$\frac{[\Gamma(\frac{1}{2})]^n}{\Gamma\left(\frac{n}{2}\right)}\int_{-1}^1 F(k\xi_1)(1-\xi_1^2)^{\frac{n}{2}-1}\,d\xi_1.$$

20. If in Ex. 18 and Ex. 19 the function $F(k\xi_1)$ is unity, show that the values of the integrals are respectively

$$\frac{[\Gamma(\frac{1}{2})]^n}{\Gamma\left(\frac{n}{2}+1\right)} \text{ and } \frac{[\Gamma(\frac{1}{2})]^{n+1}}{\Gamma\left(\frac{n+1}{2}\right)}.$$

21. If the region of integration is the octant of the ellipsoid

$$x^2/a^2+y^2/b^2+z^2/c^2=1,$$

for which $x \geqq 0$, $y \geqq 0$, $z \geqq 0$ and if α, a_1, b_1, c_1 are positive, prove that

$$\iiint \frac{dx\,dy\,dz}{(\alpha+a_1x^2+b_1y^2+c_1z^2)^{5/2}}=\frac{\pi}{6}\,\frac{abc}{\alpha\sqrt{\{(\alpha+a_1a^2)(\alpha+b_1b^2)(\alpha+c_1c^2)\}}}.$$

22. If the variables are never negative and are such that the sum of their squares lies between 0 and 1 (0 and 1 included), prove that

(i) $\displaystyle\int_0^1 \left(\frac{1-x^2}{1+x^2}\right)^{\frac{1}{2}} dx=\frac{[\Gamma(\frac{1}{4})]^2}{4\sqrt{(2\pi)}}-\frac{\pi\sqrt{(2\pi)}}{[\Gamma(\frac{1}{4})]^2};$

(ii) $\displaystyle\iint\left(\frac{1-x^2-y^2}{1+x^2+y^2}\right)^{\frac{1}{2}} dx\,dy=\frac{\pi}{8}(\pi-2);$

(iii) $\displaystyle\iiint\left(\frac{1-x^2-y^2-z^2}{1+x^2+y^2+z^2}\right)^{\frac{1}{2}} dx\,dy\,dz=\frac{\sqrt{(2\pi)}}{8}\left\{\frac{4\pi^2}{[\Gamma(\frac{1}{4})]^2}-\frac{1}{6}[\Gamma(\tfrac{1}{4})]^2\right\};$

(iv) $\displaystyle\iiiint\left(\frac{1-x^2-y^2-z^2-u^2}{1+x^2+y^2+z^2+u^2}\right)^{\frac{1}{2}} dx\,dy\,dz\,du=\frac{\pi^2}{16}\left(1-\frac{\pi}{4}\right).$

23. Prove that

$$\iiint\left(\frac{a^2+x^2+y^2+z^2}{a^2-x^2-y^2-z^2}\right)^{\frac{1}{2}} dx\,dy\,dz=\frac{a^3\sqrt{\pi}}{3\sqrt{2}}\left\{[\Gamma(\tfrac{1}{4})]^2+12[\Gamma(\tfrac{3}{4})]^2\right\},$$

where the region of integration is bounded by the sphere $x^2+y^2+z^2=a^2$ (not merely that octant for which the coordinates are positive or zero).

24. If $U=\displaystyle\int_0^\infty\int_0^\infty e^{-\left(x+y+\frac{a^3}{xy}\right)} x^{\frac{1}{3}-1} y^{\frac{2}{3}-1}\,dx\,dy,\ a>0,$

prove that

$$\frac{dU}{da}=-3U,\ U=\Gamma(\tfrac{1}{3})\Gamma(\tfrac{2}{3})e^{-3a}=\frac{2\pi}{\sqrt{3}}e^{-3a}.$$

In the integral obtained by differentiation let $\xi=a^3/xy$ and change from x to ξ; the integral when thus transformed is U.

25. If $U=\displaystyle\int_0^\infty\int_0^\infty\ldots\int_0^\infty e^{-F}G\,dx_1\,dx_2\ldots dx_{n-1},$

where $$F=x_1+x_2+\ldots+x_{n-1}+\frac{a^n}{x_1x_2x_3\ldots x_{n-1}},\ a>0,$$

and $$G=x_1^{\frac{1}{n}-1}x_2^{\frac{2}{n}-1}\ldots x_{n-1}^{\frac{n-1}{n}-1},$$

prove that $$\frac{dU}{da}=-nU,\ U=Ce^{-na},$$

and $$C=\Gamma\left(\frac{1}{n}\right)\Gamma\left(\frac{2}{n}\right)\ldots\Gamma\left(\frac{n-1}{n}\right)=(2\pi)^{\frac{n-1}{2}}.n^{-\frac{1}{2}}.$$

26. If $n > 0$, the volume within that part of the surface

$$(x/a)^n + (y/b)^n + (z/c)^n = 1$$

which lies in the first octant is $\dfrac{abc}{3n^2}\dfrac{[\Gamma(1/n)]^3}{\Gamma(3/n)}$

INDEX

The numbers refer to the pages

PRINTED IN GREAT BRITAIN BY ROBERT MACLEHOSE AND CO. LTD.
THE UNIVERSITY PRESS, GLASGOW